-

The Complementation of the Verb *Warn* in Recent Centuries

Matti Kettunen
University of Tampere
School of Language, Translation and Literary Studies
English Philology
Pro Gradu Thesis
September 2013

Tampereen yliopisto
Englantilainen filologia
Kieli-, käännös-, ja kirjallisuustieteiden yksikkö

KETTUNEN, MATTI: The Complementation of the Verb *Warn* in Recent Centuries

Pro gradu –tutkielma, 81 sivua
Syyskuu 2013

Tässä korpuspohjaisessa Pro gradu –tutkielmassa tarkastellaan englannin kielen verbin *warn* komplementaatiota kirjoitetussa brittienglannissa 1710-luvulta 1990-luvulle. Tarkoituksena on ensisijaisesti selvittää mitä komplementteja kyseisen verbin kanssa esiintyy, miten niiden käyttö on muuttunut viime vuosisatojen aikana sekä missä määrin eri komplementit korreloivat tiettyjen merkitysten kanssa.

Tutkimuksen aineisto on saatu kahdesta elektronisesta korpuksesta. The Corpus of Late Modern English Texts (CLMET) tarjoaa aineiston vuosille 1710–1920 ja korpuksesta on hyödynnetty kahta eri versiota riittävän aineiston saamiseksi: vuosien 1710–1780 aineisto on saatu uudesta CLMET 3.0:sta, kun taas vuosien 1780–1850 sekä 1850–1920 aineisto saatu CLMETEV:stä (Extended Version). Nykypäivän data puolestaan on peräisin British National Corpuksesta (BNC), joka kattaa vuodet 1960–1993. BNC:stä saatu materiaali rajattiin tekstityypin perusteella kaunokirjallisuuteen, jotta sitä voitiin vertailla CLMET:stä saatuun aineistoon. Näiden lisäksi tutkimuksen aikana mielenkiintoa herättäneeseen objektittomaan rakenteeseen haettiin erillistä lisämateriaalia käyttäen lähteenä koko BNC:n kirjoitetun englannin osiota.

Tutkielma koostuu kahdesta osasta: Ensimmäinen osa esittelee käytössä olevat korpukset. Lisäksi siinä tarkastellaan ja selvennetään korpustutkimuksen ja komplementaation keskeisiä käsitteitä ja komplementaatioon liittyviä teorioita sekä muodostetaan alustava kuva tutkimuksen kohteena olevasta verbistä valikoitujen sanakirjojen ja kielioppiteosten avulla. Toisessa osassa keskitytään korpusaineiston analyysiin. Jokainen ajanjakso analysoidaan omana kokonaisuutenaan ja näitä ajanjaksoja verrataan toisiinsa. Lopuksi objektittomia rakenteita käsitellään erillisessä osiossa.

Tutkimus osoittaa, että verbin *warn* yleisimpien komplementtien frekvenssissä ja käytössä on tapahtunut selkeitä muutoksia, joista tärkeimpinä mainittakoon *to*-infinitiivin käytön vähentyminen ja *that*-sivulauseiden sekä pelkän nominilausekkeen (NP) käytön lisääntyminen. Näistä viimeksi mainittu on ensimmäistä tutkimuksen kohteena ollutta ajanjaksoa lukuun ottamatta ollut yleisin komplementti. Lisäksi *warn* kokonaisuudessaan on nykypäivänä moninkertaisesti käytetympi kuin 1700-luvulla. Esitellyistä teorioista The Complexity Principle osoittaa erityisesti myöhemmillä ajanjaksoilla vaikuttavansa verbin *warn* komplementaatioon.

Avainsanat: warn, verbi, komplementaatio, korpus, korpuslingvistiikka

Table of Contents

1 Introduction

Consider the following examples, taken from the Corpus of Late Modern English Texts and the British National Corpus:

> (1) He had forgotten to *warn* the queen of the bowl, which she drank of ... (Lamb 1807, *Tales from Shakespeare*)
>
> (2) ... and he darkly *warned* me that I should suffer if 'anything got out'. (Gosse 1907, *Father and Son*)
>
> (3) He had overheard his uncle *warning* the two maid-servants not to go up there because it was too dangerous. (B1X 59)

All three examples feature the verb *warn*. What is different in its use, however, is the type of element it is complemented by: a noun phrase (NP) followed by *of* NP in the first one, an NP followed by a *that*-clause in the second, and an NP followed by a *to*-infinitive clause in the third.

The purpose of this thesis is to examine the complementation patterns of the verb *warn* in British English from the 1700s to the present day, aiming to define which complements are possible, what their frequencies are, and how has the use of the verb in this respect changed in the last three centuries. Additionally, semantics will also be taken into account; the different possible senses of *warn* will be discussed, and sense will be related to structure in order to identify any possible correlations between the two that might have existed or exist today. In summary, my aims are to:

> (i) identify the complement patterns that have occurred with *warn* during the last three centuries and present their respective frequencies,
>
> (ii) examine what changes have possibly taken place in the distribution of the different patterns between the different eras,
>
> (iii) explore to what extent syntactical form correlates with the sense of the verb, i.e. if certain patterns favour certain senses,
>
> (iv) see whether some selected theories and principles that relate to complementation hold as regards *warn*, and
>
> (v) consider whether the findings may have broader implications for the study of complementation.

The data for the study is drawn from two electronic corpora of English: 18th and 19th century material is taken from two different versions of the Corpus of Late Modern English Texts, and for present-day English data I have used the British National Corpus. The Late Modern English era is of interest as it is arguably "the most neglected period in the history of the English language" (De Smet 2005, 69).

More generally, the reasons for conducting a study on complementation using corpus material are twofold. Firstly, corpus linguistics in its present form (with colossal electronic corpora widely available) has barely reached adulthood, and new research can be considered both welcome to the linguistic community and intriguing to the person conducting the study. Secondly, studies in this field can be of help in EFL[1] teaching. Though we already have grammars that have made use of corpus material and that can guide the learner towards the right direction with regard to verbs and their complements, "[corpora-based] frequency information can help a teacher to decide which items to emphasize" (Conrad 2000, 553). Examining a single word, then, can be considered one of the pieces in a larger puzzle, as it were, that might be helpful when seeking to exemplify the relation between a given type of words and their complement selection, thus contributing to the integration of grammar and vocabulary in teaching – which, according to Conrad (ibid.), is a possible outcome of the rise of corpus linguistics.

This thesis can be loosely divided into two parts. The first part will begin with a brief description of a corpus and corpus linguistics, followed by an introduction of the corpora used as the source of data in this study. After this, key concepts and theories of complementation will be explained, before moving on to an overview of what has been written on *warn* in some selected dictionaries and grammars of English. In the second part I will then proceed to the analysis of the data. There are altogether four sets of data from the different eras, and these will be discussed separately in chronological order. Finally, the results will be compared and conclusions drawn.

[1] English as a Foreign Language

2 On Corpora

In this chapter I will discuss the notions of corpus linguistics and a corpus, followed by an introduction of the corpora used in this study.

2.1 Corpus and corpus linguistics

According to Hunston (2002, 2), the word *corpus* can be defined as "a collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic study." Generally speaking, a corpus is always compiled for a certain purpose. More general corpora are useful when searching answers to various linguistics questions – be it grammar, lexis or pragmatics (Kennedy 1998, 3-4). More specialized corpora, in turn, can be used to investigate a very context-specific linguistic phenomenon, such as the language used by, for instance, workers in the oil industry (ibid.).

It is worth stressing that, as of late, the word *corpus* has invariably been used with reference to electronic collections in particular (Hunston 2002, 2.). Modern corpora, the largest of which are hundreds of millions of words in size, can be processed with computers, using corpus access software, enabling quick and effective searches into the data. Logically enough, c*orpus linguistics*, then, can be defined as "the empirical study of language relying on computer-assisted techniques to analyze large, principled databases of naturally occurring language" (Conrad 2000, 548). Kennedy (1998, 4) adds that a "corpus can be analysed and compared with other corpora or parts of corpora to study variation," and that "most importantly, it can be analysed distributionally to show how often particular phonological, lexical, grammatical [...] features occur, and also where they occur." Essentially, this is at the centre of the present study as well.

What is also noteworthy is that there exists a division between so called corpus-driven and corpus-based research approaches. A corpus-driven approach "makes minimal a priori assumptions regarding the linguistic features that should be employed for the corpus analysis," and therefore

aims to "document the existence of linguistic constructs that are not recognized by current linguistic theories" (Biber 2010, 160-2). Conversely, instead of discovering new linguistic features, a corpus-based approach aims to "discover the systematic patterns of use that govern the linguistic features recognized by standard linguistic theory" (ibid., 163). Furthermore, corpus-based studies often expose variation between different linguistic variants across different contexts (ibid.). It follows that this study can more naturally be categorised as corpus-based.

2.2 Possible problems of corpus data retrieval

Among the advantages of electronic corpora, Kennedy (1998, 5) lists "incredible speed, total accountability, accurate replicability, statistical reliability and the ability to handle huge amounts of data." In short, we can access and analyse vast amounts of data in a short amount of time. Unfortunately, this is not the whole truth. Obviously, a corpus is always restricted to certain kind of linguistic data, albeit many large present-day corpora include texts from various domains and, in this way, try to be representative of as large a group of language users as possible. Additionally, there are more specific pitfalls to be aware of when conducting corpus searches.

Firstly, as Ball (1994, 295) points out, even if a corpus is automatically tagged, one should abstain from assuming that a search yields perfect results. That is, if we search for *warning* used as a verb, we might still get tokens of nominal or adjectival usage, incorrectly tagged as verbs. Therefore, we should always, if even remotely possible, supplement the automated process by manually checking the results for accuracy (ibid.). Secondly, the concepts of *recall* and *precision* should be taken into account, with the latter referring to "the proportion of retrieved material that is relevant" and the former to "the proportion of relevant information that was retrieved." The aforementioned example is a case of precision, the problems wherein, according to Ball, often result in "the narrowing of the search criteria, which in turn may lead to a decrease in recall" (ibid.). In other words, if one tries to automatically exclude the irrelevant examples, there is a risk that

relevant tokens will also be lost in the process. With this study, however, the phenomenon (i.e. the complementation of a verb) under scrutiny is such that we can expect extremely high recall. Moreover, the relatively small amount of data enables manual checking, resulting in perfect precision within the retrieved data.

2.3 Normalized frequencies

When comparing frequencies that come from different sets of data – i.e. different corpora and sub-corpora – we must appreciate that the raw numbers are often not directly comparable, as different collections of texts are rarely equal in size. It is crucial to note that a set of 10 tokens drawn from a corpus of 2 million words is not equivalent to 10 tokens drawn from a corpus of, say, 5 million words. To make raw numbers comparable, a process known as normalization (or norming) is generally applied, in which the raw number of tokens is divided by the total number of words (the size of the corpus or sub-corpus the data is drawn from) and then multiplied by the basis chosen for the normalizing (cf. e.g. Biber 1998, 263-4). For example, if we use 1 million as the basis for normalizing, 15 tokens retrieved from a collection of 2.5 million words gives us the following formula:

(15 tokens / 2.5 million words) x 1 million,

which would result in a normalized frequency of 6 tokens per 1 million words.

2.4 The corpora used in this study

I will now introduce the two corpora used in this study: The Corpus of Late Modern English Texts and the British National Corpus.

2.4.1 The Corpus of Late Modern English Texts

The source for historical data in this thesis is the Corpus of Late Modern English texts (hereafter CLMET). Compiled by Hendrik De Smet, the corpus is divided into three 70-year sub-periods as

follows: 1710–1780, 1780–1850 and 1850–1920. The original version contains some 10 million words, but for this study I have chosen to use the more recent versions of the corpus: the 15 million word Extended Version (CLMETEV) and the new CLMET 3.0, which comprises some 34 million words and includes the earlier versions. In the early stages of this project, the latter of these was not yet available. However, as the CLMETEV proved insufficient in its data (only 7 tokens found) for the period of 1710–1780, and as it was recently brought to my attention that a new version has surfaced, I decided to draw the data for the first era from the CLMET 3.0 to cover all the time periods. The following table summarizes the make-up for the relevant sub-sections in the CLMET 3.0 and the CLMETEV:

Sub-period	Number of authors	Number of texts	Number of words
1710–1780 (CLMET 3.0)	51	88	10,480,431
1780–1850 (CLMETEV)	46	64	5,723,988
1850–1920 (CLMETEV)	51	80	6,251,564

Table 1. The corpus make-up of the relevant sections of the CLMET 3.0 and the CLMETEV.

The texts in the CLMETEV have been obtained from *Project Gutenberg* and the *Oxford Text Archive*, with some additional sources included in the new CLMET 3.0. When working with the corpus, the following four points should be borne in mind (adaptation of De Smet 2005, 70-2):

1. The texts included in one sub-period are written by authors born within a correspondingly restricted time-span. This decreases the homogeneity between sub-periods and prevents the work of any one author of appearing in more than one sub-period. It should thus bring forth any possible historical trends more clearly. A disadvantage is that the work of some authors might, as a result, be excluded.

2. All authors are British, which restricts dialectal variation within the data.

3. To avoid the "thwarting of the data by the idiosyncrasies of individual writers," the maximum number of words by any one author is limited to 200,000.

4. In spite of the measures taken to counteract the biased nature of the texts (in terms of register and the social status of the authors) drawn from *Project Gutenberg* and the *Oxford Text Archive*, the CLMET nevertheless remains biased to literary texts written by higher class male adults.

It follows that the corpus is "unfit for any fine grained sociolinguistic analysis" and, due to the sometimes obscure bibliographical history of the texts, "had better not be used for the study of phenomena that might lightly attract editorial interventions – for example, matters of punctuation, spelling-related issues [...]" (De Smet, 78-9). For our present purposes, however, the corpus provides suitable material to work with.

2.4.2 The British National Corpus

The present-day English data is taken from the British National Corpus (hereafter BNC), a 100-million word diachronic corpus that features both written and spoken material of British English. The vast majority of the texts fall under the category of 'Written books and periodicals,' which comprises roughly 80% of the corpus. Altogether, a total of 90% of the data in the BNC come from written sources, with the remaining 10% comprised of spoken material. This division is largely due to the fact that acquiring and transcribing spoken material is as expensive as it is time-consuming.

According to the *Reference Guide for the British National Corpus* (Burnard, 2007), the corpus spans over three decades, from 1960 to 1993. There is a very clear emphasis, however, on the period from 1985 to 1993, as more than 90% of all data comes from these years. The number of words from a single source is limited to 45,000 (ibid.).

The Imaginative Prose section, from which the data for this study has been taken, comprises some 16.5 million words and 476 texts. This section has been chosen as it bears the closest resemblance to the texts in the CLMET and can therefore be considered the most suitable one in making comparisons between the two corpora.

3 Complementation

In this chapter I will discuss some notions and theories central to complementation.

3.1 The complement vs. adjunct distinction

In order to study complementation, we must clarify for ourselves the meaning of a complement in English. This, however, is not always a straightforward matter. For a broad definition we can turn to Huang (1997, 75), who writes that complements "help complete the meaning of a sentence as required by the verb." A good rule of thumb, it includes the important notion of the verb "requiring" a certain kind of complement[2]. This is perhaps the strongest argument for the distinction between complements and *adjuncts* – another kind of element that relates closely to verbs in clauses – and is also taken up by Huddleston & Pullum (2002, 221) in their discussion of 'obligatoriness,' in which they state that complements are often, though not always, obligatory, whereas adjuncts are invariably optional (ibid.). They illustrate the point with the following examples:

(1)	a. She perused <u>the report</u>.	b. *She perused.	[obligatory complement]
(2)	a. She read <u>the report</u>.	b. She read.	[optional complement]
(3)	a. She left <u>because she was ill</u>.	b. She left.	[optional adjunct]

The case of obligatoriness versus optionality is clearly manifested in the above sentences. In (1) the omission of *the report* results in ungrammaticality, whereas *because she was ill* in (3) can be left out without affecting the grammatical outcome of the whole; (3b) is still a most acceptable sentence of English, merely omitting the extra information as to *why* she left.

We must understand, however, that, as a verb can have several meanings, we should always relate our analysis to a particular *sense* of a verb, rather than that verb in general. Herbst et al. (2004, xxxi) specify that, in addition to making the sentence ungrammatical, another result of omitting an obligatory complement can be a change in the meaning of the headword. In their

[2] In this paper we are interested in verbal complementation only, but other elements, such as adjectives and nouns, take complements too.

example *Many of the artists continue to live and work in St. Ives*, *St. Ives* is an obligatory complement if we take *live* to mean 'to reside,' but an adjunct if we read it in its other sense, 'be alive.' This example also serves to demonstrate that the grammatical form of a given constituent does not necessarily determine its function; depending on the head verb or, as in this case, the sense of the verb, the same constituent (here a prepositional phrase) can function both as a complement and an adjunct.

Turning to non-obligatory complements, we see that *the report* in (2) above is optional (omitting it will not render the sentence ungrammatical nor change the meaning of 'read'), but it is nevertheless considered a complement. As this is not a case of what the verb requires, it is useful to introduce another criterion that Huddleston & Pullum (2002, 219) call *licensing*: "The most important property of complements in clause structure is that they require the presence of an appropriate verb that licences them." The examples given are the following:

(4) a. She mentioned the letter. b. *She alluded the letter.
(5) a. She thought him unreliable. b. *She said him unreliable.

In the above sentences, *mention* and *think* license (or 'select') the elements that follow them, whereas *allude* and *say* do not. "This type of dependence between complements and their head verb is commonly known as *subcategorisation*" (ibid.). Were we to insert the NP *the report* from (2a) into, say, (5b), the nonsensical string **She said the report* would result. Without *the report* in quotation marks – or a very specific kind of intonation – we would surely not come across such a sentence or utterance. It appears that, while *the report* was optional after the head verb *read*, it nevertheless was not randomly selected. While it would be okay with other verbs of similar (semantic) class (like *peruse* in (1)), it is licensed by the head verb, and changing the verb to a different type will result in ungrammaticality. With adjuncts, however, one can change the verb to virtually anything and the adjunct remains unaffected. As Huddleston & Pullum state, "an adjunct such as *for this reason*, *at that time*, *however*, etc., is not restricted to occurrence with a particular kind of verb" (2002, 219).

To summarize, an adjunct a) can always be left out of a sentence and b) is, with regard to grammatical acceptability, equally happy with all kinds of head verbs (though some may of course sound more natural with certain verbs than with others). This is further illustrated in my own examples below. Note how (6b) and (6c), with an adverb phrase and a prepositional phrase functioning as adjuncts, show that "the form of the adjunct is not determined by the governing verb" (Herbst et el. 2004, xxiv).

(6) a. They read the letter. [complement OK, no adjunct]
b. They read the letter quickly. [complement OK, adjunct OK]
c. They read the letter for this reason. [complement OK, adjunct OK]
d. *They hinted the letter for this reason. [a bad complement, adjunct OK]
e. They ate quickly. [no complement, adjunct OK]

One more test, and the one that concludes my discussion of the distinction between complements and adjuncts, has to do with the expression *do so*. According to Lakoff and Ross (1966, II 4-6), a verb phrase with a non-stative verb can be replaced by the phrase *do so*. Consider their examples:

(7a) Harry forged a check, but Billy could never bring himself to forge a check.
(7b) Harry forged a check, but Billy could never bring himself to *do so*.

It is claimed that "*do so* replaces all of the constituents of the verb phrase and only these. Thus, elements that may occur after *do so* are outside of the verb phrase (are not constituents of the VP), and elements that cannot so occur are inside the verb phrase" (ibid.).

(8) John took a trip last Tuesday and I'm going to *do so* tomorrow.

In (8) above (from Lakoff and Ross) we see that *do so* corresponds to *took a trip*, but that *last Tuesday* is outside the verb phrase. *Tomorrow* can occur after *do so*, so it must be an adjunct rather than a part of the VP. Though Lakoff and Ross go further on the matter, the summary presented here neatly demonstrates the essential point of the *do so* test.

It should be pointed out that there exists more detailed discussion in grammars on the distinction between complements and adjuncts. Cases can be found that are beyond the scope of the theories introduced above, and, as Herbst et al. (2004, xxxiii) underline, "the boundaries [between

complements and adjuncts] are, quite obviously, subject to gradience." However, I do not see it necessary to discuss the distinction in more detail here, but rather wish to attend to any exceptional cases as, or if, they emerge in my own data.

Finally, it is also noteworthy that the term complement is sometimes used in a broader sense where it includes the subject of the sentence as well – a stance taken by Herbst et al., for instance. Huddleston & Pullum (2002, 216) differentiate between *internal* and *external* complements, where the former refers to the material that follows the verb[3], and the latter term is used for the subject of the sentence. In this study, *complement* is used to refer to the internal type only – that is, subjects are not treated as complements of the verb.

3.2 Valency theory

The above distinctions are central to what is known as valency theory, which, in essence, "is concerned with relationships between the verbal predicate and the other elements making up a predication" (Somers 1984, 508). Originally introduced by the French Lucien Tesnière (ibid., 507), the theory holds that a verb determines how many other elements are needed to form a grammatical sentence (Herbst et al. 2004, xxiv), and the number of such elements (i.e., complements) constitutes its valency (adjuncts, in turn, are not a part of a verb's valency) (ibid.). This number usually includes the subject of the sentence as well, but as previously stated, we are, on syntactical level, only interested in post-verbal complements.

In addition to examining the syntactic level, we can relate items to the head verb on a semantic level as well. This will be discussed in the following section.

[3] In the case of extractions (cf. section 3.5.2), however, the complement can be moved and therefore does not necessarily occur in post-verbal position, although the underlying semantic structure remains unaffected.

3.3 Argument structure and theta theory

The argument structure of a given predicate lists "the participants minimally involved in the activity or state expressed by the predicate" (Haegeman, 1991, 36). Like the valency of a verb, its argument structure includes the subject as well, so we are concerned with more than just the elements that are inside the VP. (Adjuncts, again, are not taken into account.) Haegeman adds that "the argument structure of the verb predicts the number of constituents needed but not necessarily their type," which means that aside from NPs, we can expect other kinds of constituents (such as clauses) to occur as well.

Verbs (or more generally, predicates) can be labelled according to the number of arguments that they require. For example, if a verb expresses activity involving two arguments, it follows that the sentence must have a matching number of constituents that "enable these arguments to be expressed" (ibid.). Such a predicate – e.g. *meet* – is known as a two-place predicate. *Give*, in turn, is a three-place predicate, as we need to express the "giver," the recipient and that what is being given. As when defining complements, mapping the argument structure of a verb should be linked to a particular sense of that verb, as different senses might have different argument structures. Depending on the context, *warn* selects one to three arguments. It is a two-place predicate in *I warned him*, but a three-place predicate in *She warned us to leave*.

In addition to giving the number of arguments a predicate takes and possibly the syntactic type of arguments it typically subcategorizes for, we can map the semantic properties these arguments project. Firstly, we can use labels such as [+human] or [-animate] to describe the distinctive semantic features of the words under scrutiny (cf. e.g. Leech 1974, 96; Radford 1988, 370). Additionally, we can examine how arguments relate to the predicate using theta roles (or semantic roles), which are "the more specific semantic relationships between verbs and their arguments" (Haegeman 1991, 41). There is much debate on the theory of theta roles, and linguists disagree as to how many, and what kinds, should be identified (ibid.). Among the most common

ones are the roles of agent (the one who initiates the action) and theme (a person or thing that undergoes or is moved by the action) (cf. e.g. Carnie 2002, 168-9; Haegeman 1991, 41-2) In *I warned him*, the predicate *warn* assigns the role of agent to the subject argument *I* and the role of theme to the object argument *him*. For any predicate, we can present a typical set of theta roles which it selects, much like we can identify a typical number of arguments it takes.

The matching of arguments and theta roles is governed by what in theta theory is known as the Theta Criterion, formulated by Haegeman (1991, 45) as follows:

> Each argument is assigned one and only one theta role. Each theta role is assigned to one and only one argument.

It follows that with every predicate in a sentence, there must be no arguments that lack a theta role and no theta roles that lack an argument. The following section discusses a factor relating to complementation where the theta criterion comes into play.

3.4 Control

Here I wish to bring up a notion that bears relevance to the complementation of *warn*, namely that of control. In short, a control construction is one where the element functioning as either the subject or object in the higher (superordinate) clause "controls" the understood subject of the lower clause. The concept of control is therefore relevant to sentences in which a non-finite (either infinitival or gerundial) subordinate clause is present. Essential to the theory of control is what is known as the Projection Principle, originally formulated by Chomsky (1986, 84), according to which "lexical structure must be represented categorically at every syntactic level." This was later extended by the assumption that subjects are obligatory in every sentence. Consider the following:

(9) John tried to do it.
(10) John tried doing it.
(11) John told Jane to do it.

In (9) – (11), the lower (non-finite) clause lacks an explicit subject. Considering (9) and (10), *John* is the subject of the higher clause and gets its semantic role from *try*. Recalling the Theta Criterion

introduced above, we understand that *John* cannot function as the subject of the lower (*to*-infinitive) clause, as this would violate the rule. Therefore, we assume an implicit, understood subject to be present (commonly represented by the symbol PRO in recent work) as we need something to fill the subject argument position and to saturate its theta role. To illustrate this further, the three sentences are reproduced in (11') – (13') below:

(9') John$_1$ tried [PRO$_2$ to do it].
(10') John$_1$ tried [PRO$_2$ doing it].
(11') John$_1$ told Jane$_0$ [PRO$_2$ to do it].

In both (9') and (10') *John* is the controller of the implicit subject PRO located in the lower clause. Unlike *John*, PRO does not get its theta role from *try*, but from *do*. Example (11') is different in that the controller is not the subject, but the object NP *Jane*. It can be concluded that in (9') and (10'), $_1$ = $_2$, that is, the NPs *John* and PRO ultimately denote the same entity. In (11') however, it is the object NP *Jane* that shares its denotation with PRO.

Warn, in its relevant sense, is an object control verb.[4] According to Rizzi's (1986, 503) formulation of what is known as Bach's Generalization, "in object control structures the object NP must be structurally represented." In other words, if the object of the matrix predicate is the controller of PRO, it cannot be omitted. It will be of interest whether the data yields examples that are relevant as regards Bach's Generalization and whether the background literature discussed later addresses the issue.

3.5 Factors relating to complementation

I will now briefly introduce some selected principles and theories that might prove useful when seeking to account for the choice of a complementation pattern in any given case. Later in the analysis section I will, where suitable, relate these principles to the examples of actual use.

[4] Not all predicates are subject/object control predicates and not all understood subjects are PROs. Another commonly acknowledged type of understood subject is known as 'trace,' which results from a process called NP movement or 'raising' (cf. e.g. Davies & Dubinsky 2004; Postal 1970). In this study, however, it suffices to restrict our discussion to control.

3.5.1 The Complexity Principle

Rohdenburg (1996, 151) has stated that "in the case of more or less explicit grammatical options the more explicit one(s) will tend to be favoured in cognitively more complex environments." This is what is known as the Complexity Principle. Explicitness is taken to mean that sentential complements are more explicit than nominal ones, finite clauses are more explicit than infinitival clauses, which in turn are more explicit than *–ing*-clauses. This idea tallies nicely with the hierarchical grouping of complements according to their "nouniness," as outlined by Ross (1973, 137). At the top of Ross's list are *that*-clauses, followed by infinitives and *–ing*-clauses. At the bottom are nouns. Rohdenburg adds to this by distinguishing between the smallest of differences, going on to state that *upon*, for example, is more explicit than *on*, as it is a 'bulkier' element (152).

The factors that make a sentence more complex and thus, allegedly, prone to favouring more explicit grammatical variants, include passive constructions, adverbial or other kinds of insertions between the head (plus the object NP) and the complement clause, and the presence of "more or less complex surface objects preceding finite and non-finite clauses" (173). *I warn <u>you</u> ...* would therefore be considered more complex a platform for the following complement than *I warn Ø*. *I warn you warmly and with the best of intentions...*, in turn, might lead us to expect that the sentence continues with a more 'explicit' *that*-clause rather than a prepositional phrase, such as *of* NP, as there is a complex insertion involved. Also, the omission of *that* in a *that*-clause is said to become more unlikely as the complexity of a sentence increases.

Rohdenburg additionally notes a tendency of treating explicit variants more formal than less explicit ones. Therefore, the Complexity Principle can be expected to conflict with stylistic and semantic tendencies (152). It will be of great interest to see how the Complexity Principle holds in my own data.

3.5.2 Extractions

According to the Extraction Principle, formulated by Vosberg (2003, 308), "in the case of infinitival or gerundial complement options, the infinitive will tend to be favoured in environments where a complement of the subordinate clause is extracted [...] from its original position and crosses clause boundaries." Consider the following (my own) examples:

(12a) It was the bridge (that) I warned them to avoid.
(12b) It was the bridge (that) I warned them of.

In (12a), the Extraction Principle can be applied, as the complement of *avoid,* which is subordinate in relation to *warn*, has been extracted by way of clefting (cf. Vosberg 2003, 307). In (12b) there is no subordinate clause after *warn*, so we cannot apply the principle as such. It is nonetheless an example of extraction – here within a sentence.

We do not expect *–ing*-clauses and *to*-infinitives to compete in the same environments in the data (but cf. 3.5.4), for, as our intuition and knowledge of English tells us, *to*-infinitives have a very distinctive meaning when used with *warn* (cf. 4.2; 4.3; 4.7). Thus, the Extraction Principle per se might not be a major point of interest for us. However, extractions in general, be it with or without a subordinate clause after *warn*, surely add to the overall complexity of a sentence and are therefore relevant with respect to the Complexity Principle introduced in 3.5.1.

3.5.3 Horror aequi

According to the *horror aequi* principle outlined, among many others, by Rohdenburg (2003, 236), there is a "widespread (and presumably universal) tendency to avoid the use of formally (near-) identical and (near-) adjacent (non-coordinate) grammatical elements or structures." *I am warning you of going there,* for example, would violate the *horror aequi* principle, as there are two (near-) adjacent *–ing* forms used in the sentence. In similar vein, *It was brave to warn them to avoid it* is to be considered a violation for the same kind of reason. We shall keep our eyes open for any corpus examples that agree with or violate the principle.

3.5.4 Bolinger's Principle and further notes

Bolinger (1968, 127) has stated that "a difference in syntactic form always spells a difference in meaning." This is indeed something that I wish to explore in the present study. However, it must be pointed out that Bolinger was mainly referring to rival infinitival and gerundial constructions with his statement. The rivalry between (*to*-)infinitives and *–ing*-clauses is "a widely recognized phenomenon in the grammatical system of the English language [...]" (Vosberg 2009, 212). The use of *–ing* forms has increased at the expense of the *to*-infinitive, which has led to the term 'Great Complement Shift' (ibid., 213). Furthermore, theories have been put forward whereby *to*-infinitives and *–ing*-clauses are said to carry meaning in their form per se (cf. e.g. Allerton 1988, 11-23; Goldberg 1995, 1-3). That is, in sentences that are otherwise identical but differ only in their choice of the complement (minimal pairs where the choice of the infinitive vs. the gerund is the differentiating factor), differences in, for instance, temporality or potentiality (i.e. if the situation is realized or not) can be hypothesized.

The comparison of *to*-infinitives and *–ing*-clauses in this respect has been a prolific area of study in the field of complementation. It is for this reason that I chose to introduce the above notions, although, as already mentioned, *to*-infinitives and *–ing*-clauses are not expected to be rivals in the same circumstances, assuming that we adopt the strict interpretation of the Great Complement Shift – that is, only consider directly linked gerunds (i.e., with no intervening preposition) as rivals to the *to*-infinitive. Occasionally, however, prepositional gerunds are included in the discussion as well, a stance taken, for example, by Leech et al. (2009, 203). In their view, constructions such as *They warned me against going* are to be considered rivals with *They warned me not to go*. Though it is unlikely to prove pivotal to the predicate under discussion, we can keep this latter interpretation in mind when examining the data. Regardless of the interpretation of the Great Complement Shift, Bolinger's Principle, if we take it with a wider reference, nevertheless remains a very useful notion for us to be aware of.

4 *Warn* in selected dictionaries and grammars

This chapter focuses on the treatment of *warn* in some selected dictionaries and grammars of English. Attention will be paid both to the senses of the verb and the complementation patterns found.

4.1 Etymology

As aptly put by Jackson (2002, 126), "the etymological information is probably the hardest of all the parts of a dictionary entry to decode [...]" In order to shed some light into the roots of *warn*, I have consulted two dictionaries: The *Oxford English Dictionary* (hereafter *OED*) and the *Chambers Dictionary of Etymology* (*CDE*). The *OED* lists three separate entries for *warn*, only one of which is relevant as concerns the recent centuries, and, therefore, this study. Both dictionaries agree that this modern use of *warn,* first recorded between 1000 and 1200, stems from the Old English *warnian* (with many alternative spellings), "to warn, take heed" (*OED* s.v. warn, *v.*[1]; *CDE* s.v. warn, *v.*). When a word has its origins in Old English, cognates in other Germanic languages are typical (Jackson 2002, 119). The Old English form *warnian* is a case in point, as it is said to be cognate with the Middle Low German *warnen*, "to warn, inform," Flemish *waernen*, "to warn, put on one's guard," and Old High German *warnôn*, *warnên,* "to warn, refl. to provide oneself, to take precautions" (*OED* s.v.). These are all traced back to the Germanic **waranōjan* and the form **war-,* "to be cautious" (ibid.).

Of the two senses mentioned under this entry on *warn* in the *OED*, the first one, "to take heed, be on one's guard, beware" became extinct after the Old English period and is therefore of little interest to us. The second one will be discussed in the following section.

4.2 *Warn* in the *Oxford English Dictionary*

After excluding the obsolete sense of *warn*, we are left with one main sense of the verb (sense II in the *OED* entry), glossed as "to make aware, to put on one's guard," which is divided further into several sub-senses. The following table summarizes the relevant of these sub-senses in the *OED*, slightly edited where it was deemed suitable. Only the senses recorded from the 1700s onwards are present, as only these are relevant for this study. A dagger (†) signifies that a sense, or a particular use of it, is obsolete as of today.

II. To make aware, to put on one's guard		
Sub-sense	**Example(s)**	**Complement(s)**
2.a. *trans*. To give timely notice to (a person) of impending danger or misfortune [...] *to warn off:* to keep away (from danger) by timely notice.	**1794**: They say it often comes to warn people of their death. (RADCLIFFE *The Mysteries of Udolpho*) **1860:** [...] And fearing waved my arm to warn them off. (TENNYSON *Sea Dreams*)	[– NP *of* NP] [– NP *off*]
3. To put (a person) on his guard, to caution *against* some person or thing as dangerous. [† with *of* or *from*]	**1809**: This will [...] warn anatomists and surgeons, against a hasty or superficial dissection of a dead body. (*Medical Journal*) **1860**: Ah love, there sure lives in man and beast Something divine to warn them of their foes. (TENNYSON, *Sea Dreams*)	[– NP *against* NP] [– NP *of* NP]
4. a, b, d. To give (a person) cautionary notice or advice with regard to actions or conduct; to caution against neglect of duty or against wrong or mistaken action or belief. [† with *from*)	**1719**: I must only warn you, that you do not charge your Coins with more uses than they can bear. (ADDISON *Dialogues Medals*) **1756**: [...] I was warned from expecting profit, by two different sets of people. (*Museum Rusticum*) **1807**: A perfect Woman; nobly plann'd, To warn, to comfort, and command. (Wordsworth, *Poems*) **1852**: [He] marshalled the village boys,... domineering over them with a fine imperious spirit that made […] his mother fondly warn him. (THACKERAY *Henry Esmond)* **1852**: […] I warn you not to do so; I warn you to read what I have written.	[– NP *that*-clause] [– NP *from –ing*-clause) [– Ø] [– NP] [–NP *to*-inf.]

	(COLLINS *Basil*) **1852**: Be warned, therefore, against seeking a false hope in the belief that my faculties are shaken. (COLLINS *Basil*)	[– NP *against -ing*-clause]
5. a, d. To inform, notify. Now only in restricted use, to notify of something requiring attention, something actual.	**1880**: The clock warned them it was time to get ready (PARR *Adam & Eve*) **1848**: The mourning being ready, and Sir Pitt Crawley warned of their arrival, Colonel Crawley and his wife took a couple of […] (THACKERAY *Vanity Fair*)	[– NP *that*-clause] [– NP *of* NP]
5. e. *absol.* or *intr.* Of a clock: To make the clicking or whirring noise which indicates that it is about to strike; to 'give warning'. *dial.*	**1894**: Every time the clock warned to strike, she felt one hour nearer her doom. (T.H. Hall Caine, *Manxman III*)	[– *to*-inf.]
6. a. To notify of something commanded; to order under penalties.	**1815**: His royal summons warn'd the land, That all who own'd their King's command Should […] (SCOTT *Lord of Isles*)	[– NP *that*-clause]
6.c. To notify (a person) to go *from*, *out of* (a place), *away*, *thence*.	**1853**: So having warned him out of London, I made an [...] (DICKENS *Bleak House*) **1868**: She never saw Laurie mount guard in the hall, to warn the servants away. (ALCOTT *Little Women*)	[– NP *out of* NP] [– NP *away*]
6.d. *to warn off*: to notify (a person) to keep at distance. Also *fig.*	**1853**: 'Pray, Mr. Rouncewell,' says my Lady, warning Sir Leicester off with the slightest gesture of her pretty hand […] (DICKENS *Bleak House*)	[– NP *off*]
6.e. To give notice to (a person) to keep *off* (private ground). Also *fig*.	**1848**: Can the fact […] give him the right to warn all others off the ground? (*Athenæum*) **1872**: All merchants being warned off from Indian commerce as poachers from a preserve. (YEATS *Growth Commerce*)	[– NP *off* NP] [– NP *off from* NP]
6.f. *Horse Racing*. to warn off the course : To prohibit (a defaulter against the laws of the Jockey Club) from riding or running horses at meetings under its jurisdiction.	**1845**: Samuel Rogers and John Braham were warned off the Course and exercising ground at Newmarket. (*Racing Cal.*)	[–NP *off* NP] (*the course*)
7. a. To summon (a person *to* a duty, place, etc.). In later use chiefly, To summon officially; to command the attendance of. Now only *Mil.*	**1809**: The constables are required to summon or as it is said to warn all the freemen to meet together yearly. (KENDALL *Trav. Northern Parts…*) **1815**: Brother, for little space, farewell! To other duties warns the bell. (Scott, *Lord of Isles*)	[– NP *to*-inf.] [– *to* NP]
7.b. (†) To call, give notice of (a meeting). *Obs.*	**1793**: The Clerk has not inserted [in the record] that the proprietor's meeting was regularly warned. (CHIPMAN *Rep. & Diss.*)	[– NP]

8. (†)To give (a person) notice to leave his employment or tenancy. Also to *warn out. Obs.*	**1850**: We're teetotally ruined… Warned out by the landlady […] (*Bentley's Misc.*)	[– NP *out*]

Table 2: The relevant senses of *warn* in the *OED*.

As can be seen, the *OED* provides a great amount of detail and lists numerous sub-senses for *warn*. On occasion, the only motive for making the distinction between two sub-senses appeared to be the difference in complementation (with the sense remaining intact), and I have combined some of these into one (as in 4 a, b, d). It should also be noted that the complement(s) column in the above table does not necessarily list all the complement patterns that can be found in the *OED* for a particular sub-sense (though quite a few are included) but merely aims to indicate the patterns in the selected examples. However, every complement pattern for *warn* found in the *OED* is featured in the table at least once.

As an additional note worth making, it is debatable whether *off* in the phrasal *warn off* construction found in 2a and 6d–f should be considered a part of the complementation pattern or whether it should be thought of as being more closely connected to the matrix verb, restricting the use of the term *complement* to other elements. To complicate things further, *off* allows inserted material between itself and the (head element of the) matrix verb and is therefore not always met strictly adjacent to the head. This matter shall be addressed later.

As exhaustive as it is, the *OED* entry does not present the clearest view (nor is it meant to do so) of the senses that might be of most interest to us. For this purpose it should prove useful to turn to a more concise dictionary for clarification.

4.3 *Warn* in the *Oxford Advanced Learner's Dictionary*

As regards the senses and patterns, I have chosen the *Oxford Advanced Learner's Dictionary* (hereafter *OALD*) to supplement the *OED*, for it features a useful and accessible entry on *warn*. Three main senses for the verb are listed, all of which are reprinted below. I have included one of the *OALD*'s examples for each and indicated its complementation pattern, followed by a list of any

additional complements mentioned that were not found in the *OED,* thus pointing out any new patterns for us to note.

> 1. to tell somebody about something, especially something dangerous or unpleasant that is likely to happen, so that they can avoid it: *I had been warned what to expect*. [– NP *wh*-clause]

New patterns mentioned here (with some modifications made to conform to the kind of labelling I use in this paper) are: [– NP *of –ing-clause*], [– *of* *], [– *that*-clause], [– NP *wh*-clause], [– *against* *], [– NP *about* *], [– *about* *] [– NP + (speech)] and [– + (speech)]. I have inserted asterisks in some of the constructions to denote what is labelled by the *OALD* as 'someone/something.' It is thus implied that it need not be an NP, but an *–ing*-clause, for instance, is possible in the place of the asterisk as well – and because of this possibility I have included them here; no *–ing*-clauses were found in the *OED* in these particular frames. [1]

> 2. to strongly advise somebody to do or not to do something in order to avoid danger or punishment: *He warned Billy to keep away from his daughter*. [– NP *to*-inf.]

Advise is listed as a synonym here, and the non-*OED* complements mentioned are: [– NP *about* *], [– *about* *], [– NP *against* *] and [– *against* *]. Unlike with sense 1, here an *–ing*-clause example is included: *The guidebook warns against walking alone at night*. [– *against –ing*-clause]

> 3. (in sport, etc.) to give somebody an official warning after they have broken a rule (warn somebody for something): *The referee warned him for dangerous play*. [– NP *for* NP]

The above example illustrates the only complement pattern mentioned for this rather special sense, which we do not expect to find in the data.

In addition, '*warn* somebody *off* (something)' [italics added] is set apart from the rest of the entry as a phrasal verb with two senses, which, for ease of reference, I have labelled P1 and P2:

> P1. to tell somebody to leave or stay away from a place or person, especially in a threatening way: *The farmer warned us off his land when we tried to camp there.* [– NP *off* NP]

[1] It must be noted that the *OALD* groups the prepositions *about* and *against* together, and they do not explicate whether or not both prepositions are possible with the same constructions. Furthermore, their Sense 1 does not provide any examples of *about/against –ing*-clause patterns (as they do with their Sense 2), so the implicature that Sense 1 occurs with *about/against –ing*-clause should be treated with caution.

P2. to advise somebody not to do something or to stop doing something: She wanted to ask him about it but the look in his eyes warned her off. [– NP *off*]

For the latter sense, an *–ing*-clause complement is illustrated in their second example: *We were warned off buying the house*. We should bear in mind the point previously made about the uncertain analysis of these kinds of examples, and I shall return to this point in the analysis of the data.

Given its smaller compass and educative function, the *OALD* presents, as expected, a truncated version of *warn* with regard to its senses but does prove helpful as it introduces more patterns. That the *OED* does not list the *OALD* sense 3 might be explained by the educative nature of the latter; the sense itself is not different, it is merely a special type of use that is separated. Nevertheless, as the *OED* seems to make distinctions (between sub-senses) based on more subtle differences than this, and with very context-specific uses, I find it rather surprising that it fails to mention the sports-related sense brought up by the *OALD* – taking into account that I consulted the online version which even includes more recent 'draft additions.'

4.4 Simplified senses

As for the correspondence between the senses in the two dictionaries, it appears that the *OED*'s senses 2a, 3, 6a, and with the notion of 'danger' omitted, also 5, correspond to sense 1 in the *OALD*. The 'advise' sense, represented both by senses 2 and P2 (the phrasal *warn off* construction) in the *OALD*, corresponds to sense 4 (a, b, d) in the *OED*. A third category can be formed by matching sense P1 in the *OALD*'s phrasal *warn off* with senses 6c, 6d and 6e the *OED*. Because the *OED*'s senses 5e, 6f, 7a, 7b and 8 each seem to be restricted to a very specific field of discourse, I have chosen to exclude these (both the sense and the construction found with it) from further discussion. This also applies to the sports-related sense, number 3, in the *OALD*. However, should these senses nevertheless emerge in the data, this will be duly reported and dealt with.

Having now presented a rather detailed and, as a result perhaps somewhat scattered, view of

the two dictionaries, I suggest a summarized version of the senses introduced above. The following table lists the three simplified senses with *OED* examples for each.

Sense	Typical examples
Sense I: To notify someone about something (often dangerous)	(a) They say it often comes to warn people of their death. (RADCLIFFE *The Mysteries of Udolpho* 1794) (b) This will [...] warn anatomists and surgeons, against a hasty or superficial dissection of a dead body. (*Medical Journal* 1809)
Sense II: To strongly advise someone in their actions or conduct, or against wrong or mistaken action or belief	(c) I warn you not to do so; I warn you to read what I have written. (COLLINS *Basil,* 1852) (d) Be warned, therefore, against seeking a false hope in the belief that my faculties are shaken. (COLLINS *Basil,* 1852)
Sense III: Sense III: To tell someone to go away or keep their distance	(e) Can the fact […] give him the right to warn all others off the ground? (*Athenæum* 1848)

Table 3: Simplified senses

There is potential overlap between Senses I and II, which is especially evident with the preposition *against*. Using both the *OED* and the *OALD* to try and tease the two apart, we can observe that, with Sense II, the emphasis is on advising someone in their conduct or beliefs, so that problems or punishment can be avoided. The anatomists and surgeons in example (b) are informed that certain actions can be dangerous, but no direct instructions are given to them (as in *I warn you not to hastily dissect a body!*). In example (d), on the other hand, the object is directly advised on what to do or believe – perhaps with the assumption that the object is prone to certain kind of behaviour. Also, guided by intuition, the *against –ing*-clause variant falls more naturally under Sense II, but with *against* NP the distinction might be less clear. Consider the following, devised example:

(1) We were warned against a boycott at the factory.

Depending on the context, the above sentence could be read as "we were warned against boycotting / not to boycott" placing the sentence under Sense II, or, possibly even as "we were warned that a boycott might take place / of a possible boycott," where Sense I would be a more suitable

interpretation.[2] We must nevertheless accept that tokens are likely to emerge where the distinction between Sense I and Sense II is susceptible to subjective interpretation.

This kind of grouping to simplified senses not only makes for ease of reference but should also prove helpful when analyzing, and aiming to draw distinctions between, the senses in actual examples of use. It is this threefold division that will be used henceforth, and according to which the senses found in the data will be grouped – keeping an eye though, as previously stated, for any of the special senses that do not conform to these simplifications.

4.5 Complementation patterns in selected previous literature

Despite its name, *A Valency Dictionary of English* by Herbst et al. (2004) is by no means a traditional dictionary of English. Rather, as its sub-title says, it is "A Corpus-Based Analysis of the Complementation Patterns of English Verbs, Nouns and Adjectives." The book, then, focuses on listing the different complement patterns that a selected set of words occur in. As might be expected, they provide a rather comprehensive list (ibid., 935-6) of the possible complements of *warn*, reproduced below, with some modifications, again, to make possible the same kind of labelling that was introduced in the previous sections: [3]

[– NP]	[– *that*-clause]	[– NP *that*-clause]
[– *about* NP]	[– NP *about* NP]	[– *about –ing*-clause]
[– NP *about –ing*-clause]	[– *against* NP]	[– NP *about wh*-clause]
[– NP *against* NP]	[– *against –ing*-clause]	[– NP *against –ing*-clause]
[– *of* NP]	[– NP *of* NP]	[– *of –ing*-clause]
[– NP *to*-inf.]	[– NP (speech)]	[– NP (sentence)]
[– NP *away from* NP]	[– NP *off*]	[– *off* NP]
[– NP *off* NP]		

[2] Four native speakers of various varieties of English were presented with these two assumed interpretations of (1). Two, both British, did not like the latter intepretation. A New Zealand English speaker thought the latter interpretation is very unlikely, but just possible. A Canadian speaker thought both interpretations are equally possible.

[3] The patterns with a *that*-clause are said to occur in at least 30% of the tokens for *warn* in the corpus data the book is based on. Simple NP-complements are not labeled for frequency in the book, as the assumption is that they are common.

Their [– NP (sentence)] pattern differs from the reporting of direct speech (represented by the pattern [– NP (speech)] only in that the speech is indirect, that is, not in quotes. The verb, in this case, is "usually separated by commas" (ibid., xvii). It should also be noted that the last four constructions on the above list are more specifically grouped under the heading 'idiomatic phrasal verbs.' Interestingly, Herbst et al. also include the pattern [– *of* NP *with* NP], with the following example:

> (2) Queensland Rail has spent heaps on remodelling Newmarket, including level crossing signs which *warn* of approaching trains with a stylised symbol of a steam engine.

The present author disagrees with the analysis, as, for reasons explained in section 3.1, it is more natural to see *with a stylized symbol of a steam engine* as an adjunct rather than a complement. Therefore, though this might be considered a borderline case, I have chosen not to include this pattern in my list. We can return to this point in the analysis of the data, if necessary.

Though the treatment in Herbst et al. is extensive indeed, the book can be considered rather specialized. For this reason, I consider it wise, for perspective, to take a look at some general, "traditional" grammars of English that, firstly, are not geared towards linguists with the intricacies of complementation in mind and that, secondly, should provide some useful discussion to supplement their listing of complements.

In their corpus-based grammar, Biber et al. (1999) mention five complement patterns for *warn*: [– *that*-clause], [– NP *that*-clause] (665), [– NP *of* NP], [– NP + *to*-infinitive] (694) and [– NP *about –ing*-clause] (742). They also mention that *warn* is found in the passive (*be warned*), but do not see it as a verb occurring in the passive to an exceptionally notable degree.

Quirk et al. (2005) agree on the patterns but are slightly more informative, adding an [– NP *wh*-clause] (1215) to the list. Unlike Biber et al., they do not separate the [– *that*-clause] and [– NP *that*-clause] patterns as such, but do mention that the NP preceding the *that*-clause is optional in the case of *warn* (1213). Also, constructions such as [– NP *of* NP] are labelled as 'indirect object + *of* + prepositional object' (1209), but their example of this, *Mary warned John of the dangers* (ibid.)

confirms that we are talking about one and the same thing.

The possibility of omitting the object NP from an [NP *that*-clause] complement is noted by Huddleston & Pullum (2002) as well, as *warn* is featured on a list of verbs that can be used intransitively, with an "unexpressed human object" (303). They illustrate with the following examples:

(3) a. I'd advise you against buying it. b. I advise against buying it.
(4) a. I must alert you to a new danger. b. *I must alert to a new danger.

In the case of *warn*, it appears both constructions are equally possible: *I warn* (*you*) *against buying it.* "We interpret the intransitives as having a human object, but it may be either general (arbitrary people), as in *That dog bites*, or specific, e.g. *you* in particular, as in a salient interpretation of *take care: it may bite* […]" (ibid.) It will be interesting to see the frequency of these kind of constructions (omitting the object NP) in my own data, especially in object control cases (cf. 3.4). It is noteworthy that no relevant examples were found in the *OED*, even though the works discussed in this section, in addition to the *OALD,* all acknowledge this kind of use.

Finally, it is noteworthy that both Quirk et al. and Huddleston & Pullum have a take on what Herbst et al. refer to as the [– NP (sentence)] type. Quirk et al (2005, 1112-1114) talk of "comment clauses" whereas Huddleston & Pullum (2002, 895-7) address the issue with their discussion on "parentheticals." Those parts of their discussion that is expected to bear relevance to the complementation of *warn* ultimately convey the same idea: according to Quirk et al., comment clauses or parentheticals are "not independent clauses, since they are defective syntactically: the verb or adjective lacks its normally obligatory complementation" (2005, 1114). In other words, in cases such as *It is time to go, he warned*, we should consider *warn* as having no complement. It is also pointed out that comment clauses can occur in the beginning, or in the middle, of the sentence as well and that they can, in the form relevant to us, be considered a sub-category of clauses that report direct speech (1114-5). Thus, the presence or absence of quotation marks is sometimes the sole indicator of the difference between a reporting clause (for direct speech) and a comment clause,

and, if the clause occurs sentence-initially, the presence or absence of a comma the sole indicator of the difference between a "regular" zero *that*-clause and a comment clause.

4.7 Senses and patterns

As one the purposes of this study is to see whether the corpus data shows a correlation between sense and structure, I think it worthwhile to group together the complementation patterns found in the dictionaries and previous literature presented earlier, according to their simplified senses which were devised in section 4.4. The table below is based on the findings made in the dictionaries and literature discussed above. With patterns where both a variant with an object NP and one without it were found in the literature, the NP is placed in parentheses, consequently representing two different patterns that differ only in terms of the presence or omission of the direct object. Though these variants will be treated separately in the discussion of the corpus data, I find that this kind of merging makes for clearer presentation in the table.

Sense	Constructions found
I	[– NP], [– (NP) *that*-cl.], [– (NP) *of* NP], [– (NP) *of –ing*-cl.], [– NP *off*], [–NP) *against* NP], [– (NP) *against –ing*-cl.], [– (NP) *about* NP], [– NP *about –ing*.cl], [– NP *about wh*-cl.], [– NP *wh*-cl.], [– (NP) (speech)], [– NP (sentence/comment)] [4]
II	[– NP], [– NP *to*-inf.], [– (NP) *about* NP], [– (NP) *about –ing*-cl.], [– (NP) *against* NP], [– (NP) *against –ing*-cl.], [– (NP) *that*-cl.], [– NP *from –ing*-cl.], [– NP *off*], [– NP *off –ing*-cl.] [– Ø]
III	[– NP *off* NP], [– NP *out of* NP], [– NP *away*], [– NP *off*], [– *off* NP] [– NP *off from* NP]

Table 3. Constructions according to their senses in the previous literature.

Taking into account the discussion in the previous sections, we can conclude that *warn* appears to be extremely prolific as regards the different kinds of complements it takes. Furthermore, though distinctions between senses might sometimes be difficult to draw, it does seem that many of

[4] But see footnote 1 in Section 4.3 for additional discussion on *about/against –ing*-clause.

the same patterns occur with more than one sense. This is especially true with senses I and II. It will be of interest to see how these patterns are manifested in the data.

5 Corpus Analysis: Historical Data – The Corpus of Late Modern English Texts

I will now turn to the analysis of the corpus data. Proceeding in chronological order, the CLMET will be discussed first, starting with the earliest sub-period, the data for which is drawn from the most extensive version, the CLMET 3.0. The second and third sub-sections are then discussed (data drawn from the CLMETEV) before moving on to the discussion of the BNC in Chapter 6.

Each section will begin with an overview stating the total number tokens retrieved from the era under discussion. The number of irrelevant tokens is also given, accompanied by illustrations of such tokens. The different complement types found are then listed in a table in descending order of frequency, with each of the four inflectional forms separated. Any tokens of zero complements, whatever their frequency, have been placed at the bottom of the table. In addition to listing the total number of tokens for each type, the table also shows the percentage of all tokens for each pattern, followed by a normalized frequency (n per 1 million).[5] This makes for easier comparison with the other sub-corpora. To aid the reader, sentential complements have been highlighted to help distinguish them from non-sentential ones.

Following the overview is a more detailed discussion of the different complement types found. A main division has been made between non-sentential and sentential complements, which will be discussed in their respective subsections. The analysis of each sub-corpus concludes with a short recap of the main findings.

5.1 *Warn* in the CLMET 3.0: 1710-1780

5.1.1 Overview of findings

The data for this section is drawn from the recent CLMET 3.0. The first sub-section, which covers the years 1710–1780, comprises 10,480 million words. Unlike the previous versions, the CLMET 3.0 is tagged for parts-of-speech, but all four inflectional forms – *warn, warned, warning,* and

[5] The percentages and NF figures have been rounded to the nearest tenth and hundredth, respectively.

warns – were searched separately to maintain consistency within the same corpus. The search yields a total of 311 tokens from which irrelevant tokens were manually rooted out. Apart from two exact replicas with *warned*, obviously generated by a technical issue, all of these were nominal or adjectival uses of *warning*:

(1) The justice has given him fair *warning* to take care of himself. (Smollett 1751, *The Adventures of Peregrine Pickle*)

(2) … he could not leave them without *warning* and remonstrance. (Johnson 1759, *Rasselas, Prince of Abyssinia*)

(3) We are therefore called upon, as it were by a superior *warning* voice. (Burke 1775, *On Conciliation with America*)

In (1) we are, without a doubt, dealing with a nominal use of *warning*. With (2), one could perhaps argue that there is an implied, unexpressed object – someone who is being warned. However, according to the *OALD*, among other dictionaries, *warning* is indeed a noun in this very frame. The construction *without warning* is therefore considered irrelevant as well. Example (3), in turn, serves to demonstrate the (infrequent) use of *warning* as an adjective, here functioning as one of the modifiers for the noun *voice*.

After excluding the irrelevant hits, we are left with 131 relevant tokens, which translate into a normalized frequency of 12.50 instances per million words. The distribution of the tokens is given in Table 5 below.

Construction	*warn*	*warned*	*warning*	*warns*	**Total**	**% of all tokens**	**NF/million**
[– NP *to*-inf.]	12	23	3	7	**44**	33.6 %	**4.20**
[– NP *of* NP]	9	12	3	2	**27**	20.6 %	**2.58**
[– NP]	5	15	2	1	**23**	17.6 %	**2.19**
[– NP *against* NP]	5	7	1		**13**	9.9 %	**1.24**
[– NP *that*-cl.]	3	4	1		**8**	6.1 %	**0.76**
[– NP *against –ing*-cl.]	6				**6**	4.6 %	**0.57**
[– NP *of –ing*-cl.]	1	2			**3**	2.3 %	**0.29**
[– NP *to* NP]	1			1	**2**	1.5 %	**0.19**
[– NP *from* NP]	1	1			**2**	1.5 %	**0.19**
[– *against* NP]		1			**1**	0.7 %	**0.10**
[– Ø]	2				**2**	1.5 %	**0.19**
Total	**45**	**65**	**10**	**11**	**131**		**12.50**

Table 5: The frequencies of the complements in the CLMET 3.0: 1710–1780.

This era yields a total of 10 different complement patterns, with an additional two instances of the 'zero complement' type:

(4) Yet, as it is the design of this narrative, not only to divert, but to *warn*, I must descend from visionary pleasure to mere matter of fact. (Pratt 1779, *Shenstone-Green*)

Though important to note, I do not regard this as a complement type as such, and have therefore placed it at the bottom of the table.

The three most frequent complements stand out clearly, with the [– NP *to*-inf.] pattern taking first place by comprising one third of all the tokens. Nevertheless, sentential and non-sentential patterns are distributed rather evenly, with 61 instances of the former and 68 of the latter.

As regards the different (simplified) senses (cf. section 4.4), Sense II was the most common, occurring in 65 tokens, with Sense I close behind with 63 tokens. Only 3 instances were analysed as representing Sense III. It should be noted that differentiating between senses, I and II in particular, was on occasion less straightforward and depended, to varying extent, on subjective interpretation – even after looking up the tokens in their original, broader textual context. However, as our primary focus lies on complementation patterns, the figures can be considered accurate enough to demonstrate the overall distribution between the senses, and I do not see it necessary to ponder individual examples and their semantic interpretation here.

We shall now take a closer look at the different patterns. Non-sentential patterns will be discussed first, before moving on to sentential ones.

5.1.2 Non-sentential complements

Perhaps somewhat surprisingly, the [– NP *of* NP] pattern is the most frequent among the non-sentential complements with 27 tokens. They are all analyzed as representing Sense I, and it would indeed be difficult to think otherwise as our intuition and general understanding of the language dictates.

(5) … I have been severely punished; and that is what I *warned* you of from my own dear experience. (Richardson 1748, *Clarissa*)

(6) … and I have sent Jack Hathaway to see how the land lies, and *warn* you of your danger. (Smollett 1751, *The Adventures of Peregrine Pickle*)

(7) The progress of their journey might, however, have *warned* him of the impending danger. (Gibbon 1776, *The Decline and Fall of the Roman Empire)*

It is remarkable that out of the eight extractions[6] that were found in the whole of the data, five occur with this pattern. One such example is shown in (5) above, where *that* is extracted from its original position, after *of*. With regard to the Complexity Principle, one might have expected the extractions to occur with more 'bulky' prepositions, such as *against* or *about*, but only one extraction is found with the former, and the latter does not appear in the data at all. It will be interesting to see whether extractions are distributed more evenly if *against* or *about* gain more ground in the following eras.

Another interesting find with this pattern is the kind of construction shown in (6), where the object NP is followed by a co-referential possessive pronoun. A total of six instances were found, five of which occur with the [– NP *of* NP] pattern. It is surely not the interpretation that the object is considered a danger to himself/herself, but is somehow unaware of it and needs to be warned. Rather, it appears that this is a distinctive type of use, perhaps characteristic of the era under investigation. In all but one example it is the noun *danger* in particular that follows the possessive, so it seems possible that *warn* NP [preposition] NP_{poss} *danger* has been a rather fixed phrase at the time. We shall see whether this kind of use appears later.

The second most frequent non-sentential complement is the simple NP, which appeared 23 times – 19 of which with Sense I:

(8) … for, *warned* by this adventure, the boy seldom crept out of his lurking-places… (Smollett 1751, *The Adventures of Peregrine Pickle*)

(9) … without receiving the least hint from my passions, which might have alarmed my virtue, or have *warned* my prudence. (Griffith 1764, *The Triumvirate*)

(10) The place too, *warned* my virtue… (Griffith 1764, *The Triumvirate)*

[6] Technically, passive constructions always involve extraction, as the object of the corresponding active clause is extracted and becomes the subject of the passive clause. If nothing else is extracted, however, 'simple' passives of this kind are not counted as extractions in this study.

Example (8) demonstrates how passive sentences have been analysed by transforming them into the corresponding active: someone or something (here *this adventure*) warned the object, *the boy*. Together with (7) presented earlier, this token also shows a relatively frequent (34 instances in the data) attestation of a [-animate] and, logically enough, also [-human] subject in the higher clause. What is more interesting, however, are cases where the *object* is [-human]. All in all, three such cases were found in the data, two of which were also [+abstract], shown in (9) and (10) above, where *my prudence* and *my virtue* are warned. (It is to be noted, however, that these two examples originate from the same source.) In addition to these, there were many cases where the reference of the object is not explicitly human but can be interpreted as such, for the object NP denotes an entity that is thought to consist of humans, such as *the inconsiderate*, *the undesigning part of the world* or *the old companion.*

There are also two tokens that could be seen to feature a comment clause (cf. section 4.6):

(11) Let me *warn* you as a friend, he has owned to me that he is in love; it cannot be with you for yesterday was the first time you ever met… (Walpole 1764, *The Castle of Otranto*)

We can reconstruct this sentence into "*I warn you, he has owned to me that…"* Although there is an object NP present, it seems as though a comment clause interpretation is reasonable. For now, let this be dubbed the [– NP]$_{\text{COMMENT}}$ pattern, for lack of a better term. However, recalling the limitations of the CLMET with regard to possible inconsistency in punctuation, we cannot be certain whether this is a zero *that*-clause (*that* omitted) or in fact an example of a comment clause, as, adopting a modern view with regard to punctuation, the omission of the comma would guide us towards a zero *that*-clause interpretation. More examples will hopefully surface as the following sub-corpora are analysed.

The [– NP *against* NP] construction was also fairly frequent among the data with 13 tokens, mainly occurring with Sense I (12), with four tokens with Sense II (13):

(12) … I further had the precaution to *warn* my Child against the danger of any affection for you. (Brooke 1765, *The Fool of Quality*)

(13) … and to *warn* them against a refusal that would be severely punished as a criminal disobedience… (Gibbon 1776, The *Decline and Fall of the Roman Empire*)

The object can be warned against a thing or a person but with Sense II the complements of the preposition *against* are more specifically NPs which refer to actions that could be converted into *–ing*-clauses, such as *the refusal* in (13) above. The subject warns the object, *them*, against *refusing*.

The less frequent non-sentential patterns include [–NP *to* NP] and [– NP *from* NP] with two tokens each, accompanied by a single occurrence of the objectless [– *against* NP]:

(14) That pow'r that kindly spreads the clouds, a signal of impending show'rs, to *warn* the wand'ring linnet to the shade… (Johnson 1749, *Irene*)

(15) Thus heaven turns evil into good; and by permitting sin, *warns* men to virtue. (Moore 1753, *The Gamester*)

(16) The view of the Turkish banners *warned* him from the hostile coast of Jaffa… (Gibbon 1776, The *Decline and Fall of the Roman Empire*)

(17) …The post … will be strictly watched … Miss Howe's Collins is remembered to be described … James Harlowe and Singleton are *warned* against. (Richardson 1748, *Clarissa*)

Examples (14) and (15) show a pattern that was not mentioned by any of the reference works consulted for this study. It should be stressed that in neither example is the *to* an infinitive marker, but a preposition. This is clear in (14) and, having consulted the *OED*, *virtue* in (15) is indeed a noun as well (though the verbal use has existed during Middle English with a slightly different sense), albeit the underlying meaning is something along the lines of "…*warns men to be virtuous*." The latter example is placed under Sense II, while the former I have analysed as Sense III, together with example (16). In both, something drives away the object; what is different is that the goal or end location is mentioned in (14), while (16) gives the source.[7]

Finally, though the example is perhaps slightly ambiguous, the preceding context in (17) leads us to interpret the token as a pattern with no object: it does make more sense that *James Harlowe* and *Singleton* are not seen as objects of the warning, but as someone who the unexpressed object is

[7] For a possible alternative analysis of (14), see Goldberg (1995, 3-4) and what she calls a 'caused motion construction.' This analysis, in short, would retain the sense of *warn* (Sense I) and derive the rest of the meaning from the construction: *to make the linnet move to the shade 'by means of warning.'* Sense III is, however, expected to be a valid category and the placement of (14) seems justifiable.

warned against. It is of great interest whether more tokens of objectless patterns emerge as we proceed on to the following eras.

We shall now shift our attention towards the sentential complements. The most frequent ones will, again, be discussed first.

5.1.3 Sentential complements

With 44 tokens, or 33.6%, the [– NP *to*-inf.] pattern is the most frequent one in the whole of the data from this period. As we would expect, the *to*-infinitive, which is perhaps the most prototypical realization of Sense II, is only found with this sense:

(18) … he had been *warned* in a vision not to offer the oblation in the name, or in the presence, of Theodosius… (Gibbon 1776, The *Decline and Fall of the Roman Empire*)

(19) … and to resist against that, seems like struggling against the Will of Heav'n, that *warns* us by these Forebodings to prevent these Accidents of Life. (Chetwood 1720, *The Voyages*...)

In 18 of the 44 tokens, the *to*-infinitive is used in the negative (*not/never* etc. *to*-inf.), as in example (18). Insertions are also frequent, as of the 15 insertions found in the data, 10 occur with the [– NP *to*-inf.] pattern. Both (18) and (19) feature an insertion, *in a vision* in the former and *by these Forebodings* in the latter. What is noteworthy is that 9 of the 10 insertions found with the *to*-infinitive occur when the *to*-infinitive is used in the negative. The sole example with an insertion in the positive is shown in (19). According to the Complexity Principle, we should expect insertions to occur more often with patterns that rank high in the scale of sententiality. Though the *to*-infinitive is only second to *that*-clauses in this respect, we might have expected more insertions to occur with the latter and less with the former. As nine of the insertions are in the negative (*not/never to*-inf.), we could, adopting the looser interpretation of the Great Complement Shift, perhaps see these cases as rivals with the [– NP *against –ing*.cl.] pattern and surmise that the *to*-infinitive has been favoured – though we must take into account the high frequency of *to*-infinitives in general.

Two cases of extractions are also found:

(20) … swallowed up in the superior gulph of miseries that my foreboding mind then *warned* me to apprehend. (Griffith 1776, *The Story of Lady Juliana Harley*)

(21) … my highest expectation is to be the wife of that free-liver, whom he so pathetically *warns* me to shun? (Richardson 1748, *Clarissa*)

As we are dealing with the *to*-infinitive, the extraction takes place across clause boundaries, as it is the object of the subordinate verb – not the direct object of *warn* – that is extracted. It follows that the Extraction Principle is applicable here. Keeping in mind the discussion in 3.5.2 and 3.5.4, we must be cautious so as not to draw any hasty conclusions as regards the distribution of extractions (across clause boundaries) between *to*-infinitives and *–ing*-clauses, as the *to*-infinitive has a very distinctive meaning with *warn*. However, if we again adopt the broader view of the Great Complement Shift (cf. 3.5.4), as we did in the case of insertions above, we could compare the negative *to*-infinitive constructions with the *against –ing* construction, as these seem to roughly correspond in meaning. Unfortunately, the extractions with the [– NP *to*-inf.] pattern both occur in the positive, so we must refrain from further analysis for the time being, and hope for more examples to emerge later. Nonetheless, we can still note the high frequency of extractions (two out of three) within this particular pattern, while only one is met with the less explicit [– NP *against* – *ing*-cl.].

With eight tokens, the [– NP *that*-clause] pattern was the second most frequent among the sentential complements:

(22) … and so *warning* the company that none should speak to them. (*David 1774, An Historical Account of All the Voyages Around the World*)

(23) … however I must *warn* you at the same time, Mr. Spelling, that I will not do any Violence to the Inclinations of my Child … (Brooke 1770, *The Fool of Quality*)

That-clauses were met both with Sense I and with Sense II (22). The only occurrence of an insertion with *that*-clauses is given in (23) where *at the same time* separates the direct object and the following *that*-clause. *Mr Spelling*, on the other hand, is slightly trickier, as we can think of it as being in apposition with *you* and thus a part of the direct object and not to be considered an

insertion. Whether or not this is the case (i.e., non-restrictive appositions not necessarily being insertions per se), we intuitively take it to make the structure more complex.

The two less frequent sentential patterns found are [– NP *against –ing*-cl.] with six tokens and [– NP *of –ing*-cl.] with three tokens:

(24) … the poet takes occasion to *warn* his countrymen against indulging the wild and irregular passion of love. (Cipper 1753, *The Lives of the Poets of Great Britain*...)

(25) … the twilight coming on *warned* me of returning home… (Griffith 1771, *The History of Lady Barton*)

All of the tokens with [– NP *against –ing*-cl.] occurred with Sense II, which makes it the second construction besides the [– NP *to*.inf.] pattern to occur with Sense II only.

5.1.4 Recap

The data in the first section of the CLMET yields a total of 131 tokens with a normalized frequency of 12.50 hits per million. These are comprised of 10 different patterns and two zero complements. The most frequent pattern is [– NP *to*-inf.], which comprises 33.6 % of all the tokens (NF 4.20), followed by [– NP *of* NP] and [– NP]. Insertions were found predominantly with the *to*-infinitive, but the data is to be considered insufficient to draw any solid conclusions with regard to the Complexity Principle. As for the senses, all but three examples feature either Sense I or II, which are evenly distributed in the data.

In addition to the possible changes in the frequencies and distribution of patterns, we are hoping the following data to provide us with more examples that either corroborate or conflict with some of the phenomena that are of interest to us. Another set of data is needed to put into perspective and see the true relevance of the findings made in this section.

5.2 *Warn* in the CLMETEV: 1780–1850

5.2.1 Overview of findings

The data for this section is drawn from the second part of the CLMETEV, which comprises 5,724 million words. A separate search for each of the four inflectional forms returns a total of 288 tokens. As in the previous section, irrelevant tokens abound with the form *warning*, as most of these are either nominal or adjectival. One adjectival use was also met with *warned* and was excluded:

(26) ...And shall I, *warned* and instructed as I am, be as easy a pray and as wretched a dupe? (Burney 1782, *Cecilia*)

Additionally, exact replicas – obviously generated by technical issues – have been left out. With the irrelevant tokens identified and excluded, we are left with 160 relevant tokens. These are listed in Table 6 below, organized as was introduced earlier.

Construction	*warn*	*warned*	*warning*	*warns*	**Total**	**% of all tokens**	**NF/million**
[– NP]	21	15	1	2	**38**	23.8 %	**6.64**
[– NP *of* NP]	16	13	4	2	**35**	21.9 %	**6.11**
[– NP *to*-inf.]	6	14	8	2	**30**	18.8 %	**5.24**
[– NP *that*-clause]	6	14	3	1	**25**	15.6 %	**4.37**
[– NP *against* NP]	1	7	1		**9**	5.6 %	**1.57**
[– NP *wh*-clause]	5		1		**7**	4.4 %	**1.22**
[– NP *against –ing*-clause]	3	1			**4**	2.5 %	**0.70**
[– NP *about –ing*-clause]		1			**1**	0.06 %	**0.17**
[– NP *from* NP]	1				**1**	0.06 %	**0.17**
[– NP *away from* NP]	1				**1**	0.06 %	**0.17**
[– NP *back to* NP]			1		**1**	0.06 %	**0.17**
[– NP *off*]	1				**1**	0.06 %	**0.17**
[– NP *off* NP]		1			**1**	0.06 %	**0.17**
[– NP *by* NP]		1			**1**	0.06 %	**0.17**
[– *of* NP]	1				**1**	0.06 %	**0.17**
[– *that*-clause]		1			**1**	0.06 %	**0.17**
[– Ø]	4	1			**5**	3.1 %	**0.87**
Total	**65**	**69**	**19**	**7**	**160**		**27.95**

Table 6: The frequencies of the complements in the CLMETEV: 1780–1850.

What immediately stands out is that the overall frequency of *warn* has increased dramatically. A normalized frequency of 27.95 per million as opposed to the 12.50 of the previous era shows that

warn is now more than twice as common. Not surprisingly, this has also increased the variation in the patterns, as we now have 16 different patterns instead of the 10 attested in the first period. The most common patterns are again easily distinguishable, as the top three comprises two thirds of all the tokens. What is striking, though, is that the order of the first three patterns has been reversed: the *to*-infinitive is now down to third place and the simple NP is at the top. This is also reflected in a more uneven distribution between sentential and non-sentential complement types: 92 tokens are now found with a non-sentential pattern, whereas 63 feature a sentential one. An additional five tokens have a zero complement.

The increase in the number of different patterns is largely due to the different kinds of prepositional and phrasal combinations that yield only one token each. There now seem to be cases of the phrasal verb use of *warn* – something that did not appear earlier – and we shall take a closer look at these shortly.

Logically enough, the increase of the simple NP complement and the decrease of the *to*-infinitive also bring about a difference in the distribution of the different senses. Sense I is now clearly more common than Sense II, with 104 occurrences of the former and 52 of the latter. Sense III is marginal with four tokens only.

5.2.2 Non-sentential complements

The most frequent complement in the data is, as mentioned, the simple NP, which now amounts to 23.8% of all tokens with 38 hits:

(27) ... 'I came to *warn* thee,' answered the sepulchar voice... (Bulwer-Lytton 1834, *The Last Days of Pompeii*)

(28) The parents *warned* the youth in the tenderest manner; but advice and prudent counsels were to him so loathsome... (Beckford 1783, *Dreams, Waking Thoughts, and Incidents*)

The [– NP] pattern is predominantly found with Sense I (example (27)) with six instances of Sense II, one of which is reproduced in (28) above. The subject is occasionally [-human] or [-animate] but

the object is always [+human], which, apart from one token, also holds true for all the other patterns (with an object) in the data.

As in the data from the previous era, the [– NP *of* NP] pattern holds the second place with a similar percentage. 35 tokens are found:

(29) The old hands *warned* the less experienced of their danger (Thackeray 1848, *Vanity Fair*)

(30) Of this I was *warned*, directly, by Baron Von Humboldt, the Prussian Plenipotentiary... (Coleridge 1817, *Biographia Literaria*)

Example (29) serves to demonstrate an object NP that is not explicitly [+human], but can be interpreted as such. Furthermore, it shows a pattern that we already made note of in the previous set of data: the object NP is followed by a possessive pronoun that is co-referential with the object, complemented by the noun *danger*. Coincidentally, six instances were again found, though taking into account the bigger size of the previous sub-corpus, the construction now appears more common. As for extractions, which were notable with the [– NP *of* NP] pattern in the previous sub-corpus, it is worth noting that only one case was found, seen in (30) above, where the preposition *of* and the object *this* are extracted.

One token of an objectless variant of this pattern was also found:

(31) ...to disseminate useful knowledge, to *warn* of danger, or to perform that most difficulty but most important of all duties... (Ellis 1839, *The Women of England*)

This sole token of the [– *of* NP] pattern was one of only two occurrences of an objectless pattern in the data.

The third notable non-sentential pattern is the [– NP *against* NP], which was found in nine tokens, both with Sense I and with Sense II:

(32) ...Mrs Graham, let me *warn* you in good time against the error ... of taking that boy's education upon yourself. (Brontë 1848, *The Tenant of Wildfell Hall*)

(33) Those scruples and that refinement against which he *warned* her, she herself thought might be overstrained... (Burney 1782, *Cecilia*)

(34) It determined me to do the very thing against which I was thus solemnly *warned*, and fly from my patron's house. (Godwin 1794, *The Adventures of Caleb Williams*)

The embedding of *in good time* in (32) is the only case of insertion with non-sentential complements in the data. Out of the five extractions found, two are met with [– NP *against* NP] and are given in (33) and (34). Though the frequencies are small, it is noteworthy that three out of the four insertions or extractions found with non-sentential complements occur with this pattern, featuring a rather 'bulky' preposition. Again, in some of the cases the idea behind the sentence is such that an *against –ing* construction could be possible, but perhaps the structure of the sentence more naturally takes an NP. Cases such as these (e.g. (32) and (34) above) I have analysed as featuring Sense II.

The rest of the non-sentential complements found in the data were highly marginal, having only one token each. One of the more interesting cases is the following:

(35) …"If you regain your position with him, of which you seem so confident, do not consider yourself secure – not even when you are crowned queen – but be *warned* by Catherine of Arragon."
- "Catherine has not the art to retain him," said Anne. "Henry will never divorce me." (Ainsworth 1843, *Windsor Castle*)

At first glance, it might appear that *by* introduces the agent of the passive clause (as it does in all the other tokens in which *by* was met). However, this is not the case here as we can see by looking at the additional context provided. It is not Catherine of Arragon who is doing the warning, but rather the speaker, who is warning Anne, the addressee, *of* Catherine of Arragon. It thus appears that *by* is being used here in exactly the same way as *of* in a [– NP *of* NP] construction.

Another peculiar pattern was [– NP *back to* NP]:

(36) …the highest aim of the writer does not extend beyond the act of *warning* the women of England back to their domestic duties… (Ellis 1839, *The Women of England*)

This reminds us of the [– NP *to* NP] pattern found in the first set of data. Furthermore, it can be considered an opposite to the [– NP *away from*] pattern discussed later.

Also present in the data were two instances of the phrasal *warn off* construction. I have included both of them here:

(37) …they seemed to *warn* off intruders on "their ancient solitary reign." (Cottle 1847, *Reminiscences of Samuel Taylor Coleridge and Robert Southey*)

(38) Some very bad faces have been *warned* off the premises... (Byron 1810-3, *Letters 1810-1813*)

With (37), I see more grounds for arguing that *warn off* should be treated as one unit, complemented here by an NP. The structure would then be [– NP] (where '–' represents *warn off*). This analysis holds even if *off* is placed after the noun *intruders*. In the case of (38), it is useful to first make the sentence into an active: *[the subject] warned some very bad faces off the premises*. It is different from (37) in that it is not *the premises* (cf. *intruders*) that is warned off, but *some very bad faces*. Let us now try to change the place of *off* in (38), as can be done in (37): **[the subject] warned off some very bad faces the premises*. The sentence is ill-formed. It follows that, here, I see a case for labelling this construction as [– NP *off* NP], treating *off* as a complement of *warn*. The *off* in (37) is a particle, whereas in (38) it is a preposition with the meaning 'away from.'

The following two examples are similar to (37) above:

(39) Secondly, I *warn* all others from the attempt to deviate from the ordinary mode of publishing a work by trade. (Coleridge 1817, *Biographia Literaria*)

(40) ... let me beg you to *warn* Hareton and the other away from me. (Brontë 1847, *Wuthering Heights*)

In these two examples, the latter of which is one of the few analysed as representing Sense III, *from* and *away from* ultimately convey the same idea as the prepositional *off* in (38). Perhaps the complexity of the object NP in (40) has governed the choice of the more 'bulky' *away from* instead of the less explicit *off*. In (38), too, the object NP is complex indeed, but – perhaps because of this – the sentence has been passivized, making it possible for *off* to occur adjacent to *warn*.

5.2.3 Sentential complements

Compared to the data from the first era, the most visible differences here are the drop in the frequency of the *to*-infinitive and the rise in that of *that*-clauses. Though the percentage of the [– NP *to*-infinitive] pattern is barely half of what is was before, it is nevertheless the third most common

construction in the whole of the data , and the most frequent amongst sentential complements with 30 tokens, 11 of which are in the negative. Out of the 12 insertions found in the data, six are met with this pattern:

(41) ...he...delicately turned the conversation, at the same time *warning* her, by a sidelong glance, not to recur to the subject again. (Brontë 1848, *The Tenant of Wildfell Hall*)

(42) Romeo urged Paris to leave him, and *warned* him by the fate of Tybalt, who lay buried there, not to provoke his anger... (Lamb 1807, *Tales from Shakespeare*)

The insertion in (41) is *by a sidelong glance*, specifying the means of warning. Example (42), in turn, features a particularly complex insertion: *by the fate of Tybalt, who lay buried there*. Not only is this insertion exceptional because it includes a (non-restrictive) relative clause, but it is also peculiar because of the *by* phrase that could be confused with the [– NP *by* NP] pattern found in (35) earlier. As we can see however, the *to*-infinitive pattern does eventually complete itself. The *by*-phrase here, then, is reminiscent of phrases such as *by the love of God* or *by my life*, used to emphasize the seriousness and earnestness of the utterance. Of the six insertions found with the *to*-infinitive, three are in the negative construction.

There are two extractions found, one of which is in the negative:

(43) This was the very tree connected with the wild legend of Herne the Hunter, which Captain Bouchier had *warned* him not to approach... (Ainsworth 1843, *Windsor Castle*)

Unfortunately, however, there are no extractions with the [– NP *against –ing*-cl.] pattern that we could possibly compare this construction with. The fact that there are none is of course in line with the Extraction Principle, but the frequencies are too low to have statistical significance of any kind.

Though the frequency of the *to*-infinitive has dropped noticeably by percentage, its normalized frequency remains similar to that found in the first sub-corpus (5.24 to 4.20). With *that*-clauses, however, the difference is reflected in both: the normalized frequency of 4.37 per million (of the [– NP *that*-clause] pattern) as opposed to 0.76 recorded earlier shows that there has indeed been a dramatic change in frequency, albeit the general frequency of *warn* has increased. Also, the

percentage of all tokens has gone up to 15.6% from the 6.1% recorded earlier. There are now 25 tokens of the [– NP *that*-clause] pattern:

(44) ...her son *warned* her on the threshold that if she said a word about it... he would fling it instantly away into the snow... (Disraeli 1837, *Venetia*)

(45) ...and I came to *warn* you this day that a precognition is in progress... (Hogg 1824, *Private Memoirs and* C*onfessions of a Justified* S*inner*)

Both of the above examples feature an insertion, *on the threshold* in (44) and *this day* in (45). Three insertions were found with this pattern. What is more striking, however, is the following example:

(46) And chearefull chaunticlere with his note shrill had *warned* once that Phoebus fiery carre in hast was climbing up the easterne hill... (Coleridge 1817, *Biographia Literaria*)

In this rather poetic example, the *chaunticlere* (a rooster) warns that the car is climbing up the hill. Two things are noteworthy here. Firstly, it is an example of an objectless pattern – one of only two in this set of data – the analysis being [– *that*-clause]. What is more, *once* is inserted between *warn* and the following *that*-clause. This is in conflict with the Complexity Principle, as omitting the object is more unlikely when the complexity of the sentence increases, for instance, because of inserted material.

The [– NP *wh*-clause] pattern, found in seven tokens, is interesting, as it was not found in the data from the previous era.

(46) ...on every hand there is something to *warn* him what he must not do. (Foster 1821, *An Essay on the Evils of Popular Ignorance*)

(47) ...it was a duty to warn him how people talked regarding his ways... (Brontë 1847, *Wuthering Heights*)

Of the reference works consulted for this study, only the *OALD* mentions this very pattern, albeit Herbst et al. mention a prepositional variant with *about*. The rather abrupt emergence of this pattern is surprising, and its frequency in the following sub-corpora is indeed of interest.

Of the less frequent patterns, there are four tokens of the [– NP *against* –ing-cl.] construction and one of [– NP *about –ing*-cl.]:

(48) ...I *warn* him, in the first place, against trusting in the number of names on his subscription list. (Coleridge 1817, *Biographia Literaria*)

(49) ...Tom Toady, who had *warned* Southdown about visiting such an abandoned woman, now besought to be introduced to her (Thackeray 1848, *Vanity Fair*)

The sole insertion found with the [– NP *against –ing*-cl.] pattern is given in (48) above. For comparison, three were found with the negative *to*-infinitive construction, a possible counterpart of this pattern, although the frequencies are too small for any further conclusions.

5.2.4 Recap

The analysis of the second sub-section, with data drawn from the CLMETEV, has shown that the general frequency of *warn* has increased. The previous NF of 12.50 per million has now gone up to 27.95. Though there are some new patterns and more variation, the most frequent patterns still stand far apart from the rest. The most notable difference with the patterns is that, compared to the previous set of data, the simple NP complement has now risen to the top, with the *to*-infinitive down to third place. Furthermore, there has been a notable rise in the frequency of *that*-clauses, as the percentage of all tokens is now 15.6% as opposed to the 6.1% seen earlier (with a normalized frequency of 4.37 instead of the modest 0.76 we saw the first section).

Insertions continue to be prolific with *to*-infinitives, though the distribution is slightly more balanced. As regards the different senses, the drop in the frequency of *to*-infinitives also leads to a drop in the frequency of Sense II, as Sense I is now clearly the most frequent one.

5.3 *Warn* in the CLMETEV: 1850–1920

5.3.1 Overview of findings

The third and final set of historical data is taken from the third section of the CLMETEV, comprising 6,252 million words. The corpus search yields altogether 493 tokens for *warn* and its inflectional forms. Once again, irrelevant hits have been excluded after manually going through the examples. Not surprisingly, the same reasons as before have resulted in the exclusion of tokens;

nominal and adjectival uses together with any exact replicas have been left out. In addition, another four (consecutive) tokens have been excluded, as they were all from the same source, occurring near each other, with *warn* in quotation marks:

(50) ... but the word "*warn*" was again spelt out. (Grossmith 1894, *The Diary of a Nobody*)

We are then left with 279 relevant tokens, listed in Table 7 below.

Construction	*warn*	*warned*	*warning*	*warns*	**Total**	**% of all tokens**	**NF/million**
[– NP]	44	48	5	3	**100**	35.8 %	16.00
[– NP *that*-clause]	24	33	6	2	**65**	23.3 %	10.40
[– NP *to*-inf.]	7	25	7		**39**	14.0 %	6.24
[– NP *of* NP]	7	11	4	1	**23**	8.2 %	3.68
[– NP *against* NP]	8	7	1	1	**17**	6.1 %	2.72
[– NP *off*]	1	4			**5**	1.8 %	0.80
[– NP *about* NP]	1	1	2		**4**	1.4 %	0.64
[– NP *against –ing*-cl.]	3	1			**4**	1.4 %	0.64
[– NP *of –ing*-clause]	1	2			**3**	1.1 %	0.48
[– NP *wh*-clause]	2	1			**3**	1.1 %	0.48
[– NP *off* NP]		2			**2**	0.7 %	0.32
[– NP *as to* NP]	1	1			**2**	0.7 %	0.32
[– NP *from* NP]	1				**1**	0.4 %	0.16
[– NP *from –ing*-cl.]	1				**1**	0.4 %	0.16
[– *off* NP]			1		**1**	0.4 %	0.16
[– *off* NP *from –ing*.cl.]	1				**1**	0.4 %	0.16
[– NP *off on –ing*-cl.]		1			**1**	0.4 %	0.16
[– NP *as to –ing*-cl.]		1			**1**	0.4 %	0.16
[– *of* NP]			1		**1**	0.4 %	0.16
[– Ø] (zero complement)	3	2			**5**	1.8 %	0.80
Total	**105**	**140**	**27**	**7**	**279**		**44.63**

Table 7. The frequencies of the complements in the CLMETEV: 1850–1920.

Variation in the complementation patterns continues to increase, as we now have 19 different constructions divided into 156 cases of non-sentential and 118 cases of sentential complementation. An additional five cases of the zero complement type are found, including two instances of the 'speech' type:

(51) "You'll swallow it," *warned* Judy ... (Blackwood 1915, *The Extra Day*)

The overall frequency of *warn* has continued its rapid climb, as the normalized frequency has gone up from 27.9 to 44.6 per million (a 60% increase), showing that the verb is now much more

commonplace than it appears to have been in the 18th century. Changes in the patterns can also be observed: the simple NP has strengthened its position as the most frequent complement and it now comprises 35.8% of all the tokens. Furthermore, *that*-clauses have continued to increase in frequency and are now a strong second, with the percentage of the *to*-infinitive in decline (although the normalized frequency has gone up, as *warn* is generally more frequent).

As concerns the different senses, Sense I is by far the most common, occurring in 189 tokens. Senses II and III, in turn, were found in 81 and 8 tokens respectively. In addition to the three main senses, one token was met with the rare 'summon' sense mentioned in the *OED* (sense 7a, cf. Table 2):

(52) Sixteen locomotives, sixteen engineers, and sixteen firemen would be needed … Two and one half minutes would be allowed for changing engines, three for watering, and two for coaling. "*Warn* the men, and arrange tanks and chutes accordingly …" (Kipling 1897, *Captains Courageous*)

It could be argued that the men are only informed of the situation so that they can prepare themselves for it, thus placing the token under Sense I. However, since the 'summon' sense does exist, it seems likely that here we have an example of it; it does make more sense that the men are summoned, and each are then given their own duties.

5.3.2 Non-sentential complements

The most frequent non-sentential complement in the data was the simple NP, the percentage of which has increased to 35.8% from the 23.8% it was in the previous set of data. The NP is now found in 100 tokens, with a normalized frequency of 16.00. It was also the most frequent of all complements, occurring primarily in the first sense, with an additional 17 tokens with Sense II. Four tokens were found in a reporting clause, something we could dub as the $[-\ NP]_{SPEECH}$ pattern. An example of this is found in (53), whereas (54) is the sole example in the data that can be analysed as a comment clause:

(53) "Not so easy," his uncle *warned* him... (Blackwood 1915, *The Extra Day*)

(54) … there will therefore remain very little – say 1000 pounds or 2000 pounds at the outside, as what will be actually yours … "This, let me *warn* you most seriously, is all that you must expect from me …" (Collins, 1868, *The Moonstone*)

This pattern, and indeed the verb *warn* in general, once again occurred all but exclusively with a human (or animal) object, either directly stated (pronouns, proper nouns) or clearly implied by referring to an entity that obviously consists of a group of humans: *the police*, *the public*, *the young*, *the government*, etc. Two cases were found with [– NP], though, where the reference to humans might be considered less clear:

(55) "… as they used in old times to *warn* the country-side by blawing the Stwun when the enemy was a-comin' … (Hughes 1857, *Tom Brown's School Days*)

(56) The weather door of the smoking-room had been left open to the North Atlantic fog, as the big liner rolled and lifted, whistling to *warn* the fishing-fleet. (Kipling 1897, *Captains Courageous*)

Here too, however, it can be inferred that it is the people living in the countryside (55) and the fishermen (56) that are being warned.

The second most frequent non-sentential complementation pattern with 23 tokens was [– NP *of* NP], although its frequency has decreased considerably:

(57) … and his doctor had *warned* him of the consequences that would follow his persistency in continuing to work… (Collins 1860, *The Woman in White*)

(58) *Warned* of that, he did his utmost to avoid an outbreak of discord … (Gissing 1891, *New Grub Street*)

In (58), we understand that someone, namely the subject, warned him *of* 'that,' hence the construction [– NP *of* NP]. It is also worth noting that there is now only one token of the pattern found earlier, where the object is followed by a co-referential possessive pronoun and the noun *danger*.

The pattern [– NP *against* NP] was the third noticeable non-sentential pattern, which is in keeping with earlier observations. It is manifested in 17 tokens, 11 of which exhibit Sense I (59), while the remaining six were found with Sense II (60):

(59) O Mowgli, Mowgli! Why did I not *warn* thee against the Monkey-Folk instead of breaking thy head? (Kipling 1894, *The Jungle Book*)

(60) My father used to *warn* us very solemnly against 'lip-service' by which he meant singing hymns of experience and joining in ministrations in which our hearts took no vital or personal part. (Gosse 1907, *Father and Son*)

Very solemnly in (60) is an example of an insertion, which is a rare phenomenon with non-sentential complements in the data, as only three such cases were found. One example occurred with [–NP *of* NP] and one with [–NP *about* NP]. As for the two extractions found with non-sentential complements, both occurred with this pattern, as in (61):

(61) I wonder if it be she against whom I was *warned* … (Haggard 1887, *She*)

Finally, it was only this pattern that featured the object being warned directly against a person, either by mentioning the name, as in (62), or using a personal pronoun, as in (61) and (63):

(62) Thought I would write a kind little note to Gowing and Cummings … *warning* them against Mr. Stillbrook. (Grossmith 1894, *The Diary of a Nobody*)

(63) I suppose he has *warned* you against me, hasn't he? (Caine 1897, *The Christian*)

Six cases were found that exhibit this kind of reference, with an additional two – example (59) and one where the object was warned against 'people' – that can be considered less direct, but ultimately have the same idea.

As regards the less frequent constructions, one of the more surprising findings was the [– NP *as to* NP] pattern, which has not occurred before. Two tokens surfaced:

(64) She *warned* them as to the danger of opening the sealed rooms. (Bennett 1908, *The Old Wives' Tale*)

(65) My kind landlady … did her best to *warn* and direct me as to my conduct in the house … (Linton 1885, *The Autobiography of Christopher Kirkland*)

In (64), it is perhaps easier to justify that [*as to* NP] is part of the complement rather than an adjunct, as the information it provides is essential to *warn* in this sentence. *As to* could be replaced by the more common *of* or *against* quite easily. However, selecting *of* here would be a violation of the *horror aequi* principle, which might help to account for the choice of the *as to* construction.

Example (65), on the other hand, is slightly trickier. The verb *warn* is in co-ordination with the verb *direct*, and one cannot be sure whether *warn* is then used without a complement, with a

simple NP ('you') complement, or if the construction is in fact [– NP *as to* NP], as it is with *direct*. In this thesis all examples of this kind have been analyzed by treating *warn* and whatever verbs it might occur in co-ordination with as equals, so that the complement at the end of the string of verbs is considered the complement of *warn* as well. Semantically, it does seem to make more sense that *warn* is not left stranded alone in (65) but is tied both to the object NP and the rest of the complement.

Three different kinds of manifestation of the phrasal *warn off* construction were found amongst the non-sentential complements. These were separated in Table 7 as [– NP *off*], [– *off* NP] and [– NP *off* NP]. I have included an example of each below, all with sense III:

(66) "Farmer Thompson," said Holmes, *warning* off Willum and the prong with his stick … (Hughes 1857, *Tom Brown's School Days*)

(67) As they rounded the corner of the terrace, and came in sight, I hobbled out to *warn* them off. (Collins 1868, *The Moonstone*)

(68) I *warned* him and his party off the premises. (Collins 1868, *The Moonstone*)

In (66) treating *warn off* as one unit, complemented here by an NP, seems clear. (67) is ultimately the same, but there it has been more natural to separate *off* from *warn*, as the NP is not as complex as in (66). In fact, it has been obligatory, as the object is a pronoun. It follows that this is not an example of an objectless pattern, and that these two are in fact the same construction that – were we to adopt the principle of including *off* with the matrix verb – could be represented as [– NP], where the dash represents *warn off*. The pattern in (68), however, is different in that here o*ff* again has the prepositional function, meaning 'away from.'

The objectless patterns were once again scarce, as in the whole of the data *warn* occurred only once without the object NP:

(69) … their energies had awakened to flash messages across the black dome of the sky to each other; telling and *warning* of the calamity happening in the world beneath. (Beesley 1912, *The Loss of the SS Titanic*)

Though the object is not visible in the [– *of* NP] pattern above, we take it to be implied. Here, *each other* would be a likely object, but it is not repeated – perhaps due to a *horror aequi* constraint.

5.3.3 Sentential complements

The most striking change with the sentential patterns – and in fact in the whole of the data – is the increase in the frequency of *that*-clauses. The [– NP *that*-clause] construction occurred in 65 tokens (NF 10.40 per million), with a percentage of 23.3%.

(70) For the angel of the Lord and the spirit of the Mahdi have *warned* me in a vision that the souls of the accursed Egyptians and of the miserable English shall leave... (*Churchill* 1899, *The River War*)

(71) The curtain now rises upon the last act of our little drama, for hard-hearted publishers *warn* me that a single volume must of necessity have an end. (Hughes 1857, *Tom Brown's School Days*)

Example (70) illustrates Sense I, which was found in 59 tokens. The remaining six involved sense II, which can be seen in (71). Of the 16 insertions found with sentential complements (19 in the whole of the data), 12 occurred with the pattern [– NP *that*-clause], as in example (70) with *in a vision* inserted. This is more in line with the Complexity Principle than what we saw in the previous sections, albeit *that*-clauses are now more prominent. Some very complex sentences were encountered with *that*-clauses:

(72) … she was frightened, as if *warned* by womanly instinct, which for the moment her ardour had outrun, that she had been too forward to a comparative stranger. (Hardy 1873, *A Pair of Blue Eyes*)

If we consider the corresponding active, we notice that here it is not an insertion but the overall complexity of the sentence that governs the choice of the *that*-clause. A passive construction, also present in (72), is said to be one of the factors adding to the complexity of the sentence. However, *that*-clauses were not among the most common constructions with passives in the data.

In addition, it is noteworthy that the complementizer *that* was omitted in six tokens, as in the following example:

(73) The chancellor begged me to recollect that he had *warned* me there was no romance to be expected. (Meredith 1870, *The Adventures of Harry Richmond*)

We could argue that the omission of *that* here is due to a *horror aequi* constraint; *that* is already mentioned, as *warn* itself is inside a *that*-clause.

Although its percentage of all tokens (now 14.0%) has decreased, the [– NP *to*-infinitive] construction was nevertheless the second most frequent sentential complement, and third overall, with 39 tokens (NF 6.24 per million):

(74) After depositing our purchases at the Clifton House, where the waiter *warned* us to put them under lock and key, I hoped that… (Bird 1856, *The Englishwoman in America*)

(75) Alethea *warned* her not to do this, but she persisted… (Butler 1903, *The Way of all Flesh*)

14 cases of these included a negative construction, as in (75). With the *to*-infinitive we also get three cases of extraction, all of which involve the extracted element crossing clause boundaries. One such case is exemplified below:

(76) I am on an adventure of which I understand little and was *warned* to speak of sparingly. (Brebner 1910, *The Brown Mask*)

These were the only cases of extraction found with sentential complements.

Among the less frequent patterns were [– NP *against –ing*-clause] (77), [– NP *of –ing*-clause] (78) and [– NP *wh*-clause] (79). Four tokens were found for the first of these and three each for the latter two.

(77) … she wished to *warn* him against expecting anything but a cold, offended demeanour. (Gissing 1891, *New Grub Street*)

(78) He *warned* me of growing miserly. (Meredith 1870, *The Adventures of Harry Richmond*)

(79) … Mr. Fairlie was *warned* beforehand how to receive them. (Collins 1860, *The Woman in White*)

A sentential counterpart for the *as to* pattern introduced with examples (64) and (65) (cf. section 5.3.2) is found in the [– NP *as to –ing*-clause] pattern:

(80) The doctor ... *warned* them as to keeping all spirituous liquors out of the close, and having the gates closed by nine o'clock. (Hughes 1857, *Tom Brown's School Days*)

Even though this is the only example found, it serves to confirm our analysis made earlier that the *as to* pattern is indeed a complement. It would be odd to try to (semantically) separate *warn* from the *as to –ing*-clause pattern here.

Finally, two cases of the phrasal *warn off* construction were also found:

(81) Bruno ... was careful to *warn* off the vulgar from applying the decisions of philosophy beyond its proper speculative limits. (Pater 1896, *Gaston de Latour*)

(82) The natives would still continue the search, but I *warned* them off on purchasing the land. (Baker 1855, *Eight Years' Wandering in Ceylon*)

What is interesting about these examples, both falling semantically under Sense II, is that they both introduce an additional preposition: *from* in (81) and *on* in (82). The following *–ing*-form surely requires this, and it does not change the fact that the whole string should be considered a complement. Nevertheless, *on* is a noteworthy choice in (82), as *from* seems equally possible as well – albeit one might argue that there is a small difference in meaning with the different prepositions. That, however, is a semantic issue of a depth we should not go into here. It suffices to consider the Complexity Principle; it could be argued that the complexity of the sentence – especially of the material following *from* – in (81) results in the choice of the 'bulkier' *from* instead of the less explicit *on*.

5.3.4 Recap

The analysis of the third sub-section of the CLMETEV – which concludes our discussion of historical data – has shown that the overall frequency of *warn* has continued to increase rapidly, with the normalized frequency now at 44.6 per million as opposed to 27.9 in the previous section. The [– NP] pattern is again the most common, but the most remarkable change is the rise in *that*-clauses, which are now the second most frequent complement in the data. The Complexity Principle does seem to have more relevance than before, as *that*-clauses were prominent in complex environments – at least with insertions. The *to*-infinitive has gone down in its relative percentage (albeit its normalized frequency has increased as a result of the total frequency of *warn* increasing as well).

We can also observe that most of the rarer constructions continue to emerge consistently, perhaps indicating that, despite their small frequency, they are well established. In addition, we have seen new constructions such as [– NP *as to* NP] that were not found in any of the reference works investigated. Phrasal and multi-prepositional constructions also seem slightly more prominent. As for

the senses, the distribution remains similar to that found in the previous section.

We shall now move on to the analysis of the contemporary data, with material drawn from the British National Corpus (BNC).

6 Corpus Analysis: Contemporary Data – The British National Corpus

This chapter is divided into two main sections. First, an analysis identical to those presented earlier will be conducted of the data drawn from the Imaginative Prose segment of the BNC. This 16,496 million word sub-section bears the closest resemblance to the texts in the CLMET, and should therefore provide useful reference material. Though the BNC includes the years 1960–1993, there is a heavy emphasis on the years 1985–1993, as 90% of the material originates from this era. By and large, we can say that the gap between the last section of the CLMET and the BNC is some 65 years.

The second part of this chapter then isolates one phenomenon that has roused our interest but of which little material has been found – the objectless patterns with *warn* (especially those relevant to Bach's Generalization (cf. section 3.4)). Nearly the entire BNC (the written part) will be searched for tokens of these, aiming to identify any common factors in the pattern, with special emphasis on the interpreted reference of the implicit (omitted) object.

6.1 *Warn* in the Imaginative Prose section of the BNC

6.1.1 Overview of findings

Although the BNC is a tagged corpus and a lemma search is therefore possible, all the verb forms were once again searched separately, as was done in the previous sections. The part-of-speech tagging does, however, prove useful, for one can search for verb forms only. Nevertheless, manual checking is needed, as the tagging system is, of course, not entirely foolproof.

The search returns 1288 tokens, which have then been thinned to 25% to make this into a suitable sample for comparison with the earlier data. This translates into 321 tokens, which were checked for any errors in tagging. Six such cases were found, which, it is no surprise, were all nominal or adjectival uses of *warning* incorrectly tagged as verbs. The frequencies of the 315 relevant tokens are presented in Table 8 below. There was a sharp increase in comment clauses and

the speech pattern and, as a result, I have decided to include their total frequencies in the table in the case of NPs and zero complements (S = speech; C = comment).

Construction	*warn*	*warned*	*warning*	*warns*	**Total**	**% of all tokens**	**NF/million**[8]
[– NP]	57	53	6	1	**117** *(S 12; C 21)*	37.1 %	**28.37**
[– NP *that*-clause]	19	40	5		**64**	20.3 %	**15.52**
[– NP *to*-inf.]	5	12	4		**21**	6.6 %	**5.09**
[– NP *of* NP]	5	7	4		**16**	5.1 %	**3.88**
[– NP *about* NP]	5	10			**15**	4.8 %	**3.64**
[– NP *against* NP]	2	2	3		**7**	2.2 %	**1.70**
[– *that*-clause]	1	5			**6**	1.9 %	**1.45**
[– NP *off*]	1	3	1		**5**	1.6 %	**1.21**
[– NP *off* NP]	3		1		**4**	1.3 %	**0.97**
[– NP *wh*-clause]	1	1	1		**3**	1 %	**0.73**
[– *of* NP]	1	1			**2**	0.6 %	**0.48**
[– NP *away from* NP]			2		**2**	0.6 %	**0.48**
[– NP *off from –ing*.cl]		1			**1**	0.3 %	**0.24**
[– NP *against –ing*-cl.]		1			**1**	0.3 %	**0.24**
[– NP *about –ing*-cl.]		1			**1**	0.3 %	**0.24**
[– NP *back to* NP]		1			**1**	0.3 %	**0.24**
[– *about* NP]			1		**1**	0.3 %	**0.24**
[– Ø]		47		1	**48** *(S 47; C 1)*	15.2 %	**11.64**
Total	**100**	**185**	**28**	**2**	**315**		**76.38**

Table 8. The frequencies of the complements in the Imaginative Prose section of the BNC: 1960–1993.

At a glance, much of the same tendencies are visible here as were in the previous section: the normalized frequency of *warn* has again increased (by some 71%) and the order of the top three patterns remains unchanged.The percentages of the first two are also similar, whereas the percentage (and, marginally, the NF as well) of *to*-infinitives has decreased. In conjunction with *that*-clauses it should be pointed out that the objectless [– *that*-clause] construction is now clearly more frequent (albeit still marginal compared to the top 5 patterns), which also results in a rise in the total frequency of objectless patterns.

Perhaps the most discernible difference, though, is that we now have high frequencies of both the speech pattern (i.e., *warn* in the reporting clause for direct speech) and comment clauses. This

[8] As we have taken a selection of 25%, the total word count is now 25% of 16,496 million words, that is, 4.124 million.

also explains the seemingly high number of the zero-complement type, which is now prominent alongside the 97 occurrences of sentential and 170 of non-sentential patterns. Both the speech type and comment clauses occur rather regularly with the NP complement. With the zero use, however, all but one are of the speech type, with only one comment clause (BNC reference codes given in parentheses):

(1) 'Mind his paws on the glass,' Carle *warned*. (A0R 2125)

(2) Just before seven she left Jed to pick out some toys and books to be taking with him — and it's got to be a portable amount, she *warned* and went to check that she wouldn't be needed for a while. (FYY 1217)

As previously mentioned, we must keep in mind that the material in the CLMET is reportedly prone to editorial interventions, which might affect punctuation. The "true" number of comment clauses in the CLMET, then, cannot be verified. However, even if we were to include the potential cases there, it is evident that comment clauses have nevertheless increased in frequency. The speech type does not seem to be affected by this issue, as, being more established, quotation marks are expected to be correct in the CLMET. Therefore, we can surmise that the frequency of the speech type has increased to an astounding degree.

As for the senses, Sense I is, as expected, most common with 217 tokens. Senses II and III are found in 87 and 10 tokens, respectively. In addition to these, there is one occurrence of the 'horse racing' sense (sense 6f in the *OED*):

(3) As a result, one of the committee lost his temper and told Filmer he was the scum of the earth and no one in racing would sleep well until he was *warned* off which was the number one priority of the world's racing authorities. (BP9 45)

The context in this example, featuring the [– NP *off*] pattern, shows that the goal is to prohibit the object, *Filmer*, from competing.

We shall now take a more detailed look of the data, starting with non-sentential complements. On aspects and findings already familiar to us, the discussion will now be kept briefer in order to make sure that the limelight is on the new findings that are likely to be of more interest.

6.1.2 Non-sentential complements

The percentage of the NP complement is similar to that found in the previous section, but its normalized frequency has of course increased as a result of *warn* generally becoming more frequent. 117 tokens are found, out of which comment clauses are met in 21 and the speech type in 12:

(4) I *warn* you, I'm a landlubber: I've never sailed a boat in my life. (CHG 2027)

(5) Money, obtained too easily, could finance disaster, her grandfather *warned* her... (K8R 618)

(6) 'Don't antagonize him,' the Doctor *warned* Bernice. (FR0 4207)

As discussed earlier, in cases such as (4) above, it is the comma after *you* that separates this comment clause from a zero *that*-clause interpretation. However, the manifestation of this kind of construction is so consistent and frequent that we must acknowledge that this type now has a firm foothold in the complementation of *warn*. Example (5) is even clearer, as the sequence of elements in the sentence is such that a zero *that*-clause interpretation would not be even possible. In (6), in turn, the quotation marks indicate that direct speech is being reported, separating it from a comment clause. The degree to which the frequency of these two constructions has increased is nothing short of baffling, and the present author cannot offer a satisfactory explanation as to why this might have happened. One point worth making, though, is that, according to Quirk et al. (2005, 1113-4), the subject in comment clauses is usually *I*, and apart from example (2) above, this does hold true in our data. Zero *that*-clauses feature plenty of other kinds of subjects as well, so a slight difference in this respect may perhaps be proposed. Nevertheless, it can be observed that in most cases, comment clauses and zero *that*-clauses (and even the speech type) are extremely similar and ultimately convey the same underlying idea and feel. Indeed, in speech these would more often than not be inseparable, albeit a change of intonation can be used to differentiate between them.

The only example of *warn* where the object did not have a human reference is found in the following:

(7) ... but Emil and I looked at each other, and I said, 'How do we *warn* that train?' (BP9 894)

This marks the only occurrence of such an object in the whole of the data as well. The object *train* is similar to *fishing fleet* in example (56) in section 5.3.2, and although non-human, we understand that it is the driver or staff of the train is who are ultimately about to be warned.

The NP complement is predominantly found with Sense I, as only 12 were met with Sense II:

(8) Two senior Ross-shire men were ordered to walk the boundary, as they knew it, being sternly *warned* before they set off: 'Mind you, your feet are on oath.' (AS7 1798)

(9) She also had the run of the library and she could of course go out, as long, he *warned* her with a sardonic look, as she didn't wander off with strange men. (FNT 1076)

In these two examples, the context provided shows that the objects are strongly advised to act in a certain way.

Other relatively prominent non-sentential patterns were [– NP *of* NP], all occurring with Sense I, followed by [– NP *about* NP] and [– NP *against* NP] with 16, 15 and 7 tokens, respectively. They occupy roughly the same positions in the table as they did in the previous section, though [– NP *about* NP] now appears slightly more common, and has now for the first time broken into the top five patterns. Out of the five extractions found in the data, two are met with [– NP *about* NP] and one with [– NP *against* NP]:

(10) In the second, I can't imagine what you have to *warn* me about. (JYE 2684)

(11)… that the man they had *warned* me against had taken it… (H9V 2314)

Though the frequencies are minute, we can note that throughout the different sets of data we have discussed, there have consistently been extractions with the [–NP *against* NP] pattern. It does seem, then, that there is some truth to this kind of "bulky" preposition providing a more likely environment for an extraction to occur. As for insertions, these were not met with non-sentential complements at all.

Though extremely marginal, objectless variants for the NP *of* NP and *against* NP constructions were also found:

(12) ...a high price, thought Joshua, for having to listen to Bill Cash *warn* of the dangers of Europe. (HNK 131)

(13) Lee got herself out into the street ... looking at posters advertising Tampax, *warning* about Aids, promising relief in the Bahamas, selling newspapers.... (CA3 1802)

Even though the objects are not made explicit in (12) and (13), we can deduce that it is people in general (whoever the warnings reach) that are being warned. It has perhaps been more convenient in these sentences to leave out the explicit object to denote its general nature. We shall discuss these kinds of cases more in section 6.2.

The phrasal *warn off* construction was once again found in [– NP *off*] and [– NP *off* NP]:

(14) Presumably, he had *warned* her *off*. (JYD 3026)

(15) 'No,' she went on coolly, 'don't waste your time trying to *warn* me *off* Rob. (HA7 2328)

As discussed earlier, there is a difference in the use of *off* with tokens such as those above (both of which fall under Sense III). In (14) *off* is a particle, making this construction more a phrasal unit than a separate verb + complement sequence, whereas in (15), *off* functions as a preposition meaning 'away from.' Though the examples of this construction as well as the [– NP *away from* NP] were few, a difference could be observed: with the former, the reference was always made to a person, that is, the subject was trying to prevent the direct object from being in contact with a certain person. With the latter construction, however, the reference was made to a thing rather than a person:

(16) You once said something about him *warning* you away from magic? (HA7 4004)

Finally, a single token of the [– NP *back to* NP] pattern was found, representing Sense III:

(17) Trent had already *warned* his charter party back to the catamaran, keeping himself between them and the big fish. (AMU 233)

This construction was only previously found in the data from the second part of the CLMET, where it also yields one token. This kind of use is reminiscent of the [– NP *away from* NP] pattern.

6.1.3 Sentential complements

With a normalized frequency of 15.52 per million and a percentage of 20.3% of all tokens, the [– NP *that*-clause] pattern is clearly the most frequent sentential pattern in the data. Sense I is featured in all but four examples which belong to Sense II. Something that catches our attention is the fact that the objectless [– *that*-clause] pattern has now become noticeable as well (NF 1.45). This pattern is found in six of the nine manifestations of an objectless construction in the data, which is rather surprising:

(18) Lady Thatcher remembered the time in the late 1970s when she had gone on television and *warned* that people were afraid of being 'swamped' by immigrants. (HNK 1102)

As in examples (12) and (13) in 6.1.2., here too we can conclude that the direct object is left implicit because we take it to be rather general in reference. Lady Thatcher warned no-one in particular, but anyone hearing her warning.

Another important point is that the insertions in the data are now exclusively found with *that*-clauses. Out of eight insertions total, seven occurred with the [– NP *that*-clause] pattern while one was found with the objectless variant.

(19) Had Jamie not *warned* her from the beginning that he was dangerous … she would have responded to his sex appeal, his hard good looks, his cool, clever wit. (JYD 813)

(20) Furiously Kate tried to wriggle herself free until she was *warned* quite unmistakably that Ace found her movements provocatively exciting. (HGM 1960)

(21) You may have succeeded with my brother, but I've *warned* you before you're wasting your time with me! (JXS 2186)

The insertions in the above are *from the beginning* in (19), *quite unmistakably* in (20) and *before* in (21). That insertions are only found with *that*-clauses is well in keeping with the Complexity Principle, as *that*-clauses are the most sentential type of complement. Example (21), however, is an exception in that the increased complexity of a sentence makes it more unlikely for the complementizer *that* to be omitted, as has been done here.

As for the two extractions that are found with sentential complements, one is met with the [–

NP *that*-clause] with *that* omitted, while the other is found with the objectless [– *that*-clause] variant, also with *that* omitted:

(22) Bamford on a Friday morning before Christmas was everything Mrs. Bissett had *warned* it would be. (CEB 2)

(23) Mercifully the Long Island Expressway, which Mrs. Meadows had *warned* them might well be one lone traffic jam, was proving to be relatively uncrowded. (JXX 1068)

These two examples, the latter of which is a particularly delicious example of extraction, are both in conflict with the Complexity Principle: the presence of an extraction makes it more unlikely for *that* to be omitted, let alone for the object to be omitted as well, as is done in (22). Both (22) and (23) are interesting counter-examples to the Complexity Principle, although we must keep in mind that it is not meant to be hard-and-fast rule. To take a closer look at these two examples, we feel that in (22) *that* could be inserted after *warned*, albeit the sentence is perhaps more fluent without it.[9] With (23), however, something is different. If we deconstruct the sentence, we get something along the lines of "*Mrs. Meadows had warned them <u>that</u> the Long Island Expressway might well be…"* The subject in the embedded clause has been extracted and is now realized as the relative pronoun *which* that has its antecedent, *The Long Island Expressway*, in the main clause. A restriction often dubbed the 'that trace effect' or the 'comp trace effect' (cf. e.g. Perlmutter 1971; Chomsky & Lasnik 1977) comes into play here: *that* cannot be followed by a trace in this kind of environment. As the subject is extracted and has left behind a trace, the insertion of *that* is blocked. These complex tokens, then, further demonstrate that on occasion other factors not explained by the Complexity Principle can dictate the choice of certain structures over others.

The third most frequent complement in the data and the second with sentential complements is the [– NP *to*-infinitive]. With a normalized frequency of 5.09 per million and a percentage of 6.6% of all tokens, it is not quite as frequent as it was in the previous set of data.

(24) I *warned* you not to take me for granted, Kathleen. (JYB 2536)

[9] Four native speakers of English were presented with sentece (22) and were asked to comment on the status of *that* after *warned*. All thought that using *that* was acceptable, but did not consider its use very idiomatic and/or likely.

Out of the 21 tokens found, eight occurred in negative constructions such as the one above. Perhaps further highlighting the decline in the prominence of *to*-infinitives is the fact that, for the first time, no extractions or insertions were found.

Turning to the less frequent patterns, the distribution is rather similar to those seen in the previous sections:

(25) If she *warned* Ace what her father expected of him he'd laugh in her face. (HGM 2664)

The [– NP *wh*-clause] construction seen in (24) above was found in three tokens and thus appears to be slightly more established than [– NP *off from –ing*-clause], [–NP *against –ing*-clause] and [– NP *about –ing*-clause], which only yield one token each:

(26) …she looked him over again carefully for the signs of his servitude, ignoring the forbidding stare that *warned* her off from probing. (K8S 34)

(27) She *warned* her child against coming to a bad end. (HGF 3035)

(28) As you will recall, I have persistently *warned* you all about taking promotion for granted… (FR9 851)

6.1.4 Recap

It is no surprise that the overall frequency of *warn* has again increased greatly, as the normalized frequency of 76.4 per million shows. The major patterns remain the same, with the simple NP at first place, followed by [– NP *that*-clause] and [– NP *to*-infinitive]. *That*-clauses have continued to become more frequent while *to*-infinitives are slightly less common. What is more, the data now lends more substantial support to the claims made by the Complexity Principle, as all insertions are found with *that*-clauses. Some interesting counter-examples are also met, as in the case of the that trace effect.

Another interesting thing is that omitting the object now appears slightly more common, albeit the number of tokens remains too low to have any true statistical significance and the tokens

were mostly met with *that*-clauses. Omitting the object is something that we shall take a closer look at in the following section.

6.2 Omitting the object with *warn*

In this section we shall take a detailed look on the objectless patterns with *warn*. Two additional sets of data will be introduced here: In section 6.2.1, a broader search will be conducted to retrieve different kinds of cases where the object is omitted, with the focus on semantic categorization of the tokens. To supplement this, we shall conduct an additional, more restricted search in 6.2.2 to retrieve tokens relevant to Bach's Generalization only, namely *to*-infinitives and *–ing*-clauses. These two sets of data are to be considered separate from those presented earlier in this study, as they single out a special type of pattern and include tokens from different text domains.

6.2.1 Objectless patterns: a general look

6.2.1.1 Overview of findings

Based on the reference works and data discussed thus far, a search string of the following kind was formulated: {warn/V} (of|about|against|that|to). This string searches for all the forms of the verb *warn* followed by any of the parenthesized words, which, it is justifiable to assume, are likely to yield relevant tokens with the objectless pattern. The unfortunate but virtually unavoidable fact is that zero *that*-clauses cannot be retrieved.

In order to get a sufficient amount of data, the entire written part of the BNC was searched for tokens, with the exception of the 'Newspapers' text type. This leaves us with a sample of 78,491 million words. Newspapers were excluded as it is characteristic of newspaper headlines to omit words and thus produce somewhat distorted results for us – an assumption confirmed by a preliminary search. With these adjustments, the search yields 1383 hits, which were then thinned to 200 randomly selected tokens. Out of these, 34 were irrelevant due to passives (29 tokens),

headlines (3), one nominal use of *warning* and one incorrectly tagged *that*-clause where *that* was a demonstrative inside an NP:

(29) Were you never *warned* of the dangers inherent in such relationships? (A6D 1242)

(30) RASTEROPS *WARN* OF OPERATING LOSS (CP3 153)

(31) [...]'s *warning* to send the next late-comer straight to Matron. (CK0 33)

(32) Had some inner clock *warned* that young Sheikh to 'come out and be separate [...]' (CDX 2776)

The number of relevant tokens, then, is 166. The frequencies of the different patterns found are summarized in Table 9 below. The normalized frequencies are also given, but merely to put the figures into some level of perspective; we must remember that this is a separate search and the frequencies are not – nor are they meant to be – comparable to those presented earlier in this study.

Pattern	**Number of tokens**	**NF per million**
[– *that*-clause]	120	8.30
[– *of* NP]	22	1.52
[– *against* NP]	16	1.11
[– *against –ing*-clause]	6	0.41
[– *about* NP]	2	0.14
Total	**166**	**11.48**

Table 9: Frequencies of objectless patterns.

That-clauses are the predominant pattern with 120 tokens with *of* NP coming second. There were no relevant instances of the *to*-infinitival pattern. [10] As for the senses, Sense II was found in 16 tokens (in [– *against –ing*-clause], [– *against* NP] and [– *that*-clause]), while the remaining 184 tokens fall under Sense I.

I will now take a closer look at the semantics involved in the omission of the direct object, with special emphasis on the interpreted reference of the omitted NP.

6.2.1.2 Semantic categorization

A striking feature in the data is that a majority of the examples can be considered rather formal in nature and to have a connection with politics. According to the background information on the

[10] Cf. section 6.2.2.2 where the [– *to*-infinitive] pattern is discussed.

tokens obtained using the highly useful BNC interface, three of the most dominant text domains in the data (the relevant tokens) are Informative: World Affairs (69 tokens), Informative: Applied Science (44) and Informative: Social Science (17). Only 5 tokens are found under the Imaginative Prose label. As for the general medium of text, 107 out of the 166 relevant tokens were found in periodicals.

The present author claims that the reference of an omitted (implicit) object is necessarily always generic to a certain degree. There are, however, different degrees of 'genericness' that are of interest, and I have used a threefold division to sort out the tokens according to how generic or specific they are. The sorting into categories has been made after scrutinising the tokens in their expanded context, and although some cases are subject to interpretation or point of view, the division presented below should suffice to illustrate the overall distribution of the tokens.

The first category includes tokens where the reference is clearly generic, and could easily be replaced by *people* without rendering the sentence odd – albeit actually using the word might sound somewhat clumsy as opposed to leaving it out. 11 such tokens were found:

(33) A notice at the bottom of the street *warned* that it was unsuitable for motors. (GW3 146)

(34) A bell may be used to *warn* of a cyclist's presence. (GXJ 126)

(35) A failure to *warn* that a product is not suitable for a particular purpose may give rise to liability: e.g. fireworks are not suitable for indoor use. (HXV 1658)

(36) A number of late-nineteenth-century thinkers did begin to *warn* against the dangers of pollution and environmental exhaustion, but they were as yet in a minority. (G0H 752)

(37) The Police have *warned* that he should not be approached. (J1M 3122)

In these kinds of cases the context, domain and style of the text support a clearly generic interpretation of the omitted object NP. Although some kind of restriction in the reference can always be argued (a warning never reaches, nor is it intended to reach, all the people in the world) these are cases where, I argue, it is natural to think of the reference as general, unquestionably unspecified. This is perhaps most clear in examples like (34). Likewise, the police in (37) are warning people in general, whoever the warning reaches.

The second category includes tokens where the "people in general" interpretation is too vague. Reasons for this include a specialized context and/or technical style that imply that it is not expected that people in general should heed the warning given. Tokens of this kind were most frequent with 113 hits:

(38) Professional teachers in the UK, even those who are enthusiastic about Papert's ideas, *warn* of the difficulties to be overcome in curriculum development and teacher training. (B77 537)

(39) Intergraph Corp [...] *warns* that while it has not yet closed its books for the first quarter, preliminary estimates indicate that turnover will be lower than expected at between $280m and $285m, and that all geographic segments were below expectations [...] (CP9 173)

(40) The manufacturer's safety data sheet *warns* of a potentially hazardous reaction between sodium borohydride and fine dispersed heavy metals but this reaction with charcoal and solid sodium borohydride are stored and handled in such a way that the risk of contact between them is avoided. (B0M 1245)

In (38) the teachers are surely not warning the general public, but rather the professionals or officials in the field of education who make decisions pertaining to the context. Likewise, in (39), it is mostly the shareholders and people otherwise linked to the stock market that are affected by the warning. In (40), in turn, the highly specialized nature of the context is enough to limit the reference to, for example, chemists handling the product or substance in question.

In addition to the kind of examples illustrated above, the reference in this category is in fact sometimes made rather explicit, though it is not found in the object position:

(41) My friend's friend advised us on the best solicitors for a rape case, but *warned* that the final bill might be enough to buy a flat in Mayfair. (AE0 3089)

(42) In moments of crisis, Socrates was guided by the utterances of a divine voice, which invariably *warned* against any further pursuit of knowledge. (H0N 1225)

(43) [In delight, I told the therapist.] She *warned* that when things vanish as quickly as this, it was quite possible that they might come back. (ADG 256)

(44) Sir Peter Imbert, the Met Commissioner, has written to Met officers *warning* that we would take proceedings under the Official Secrets Act against police who break an internal discipline code [...] (A0K 408)

(45) He *warned* that if Slovenia went ahead with secession, EC support would be withdrawn – leading to the total economic collapse of Yugoslavia as early as September. (HL8 219)

In (41) – (44) the implied reference of the omitted object is relatively clear. It might even be argued that target of the reference is present in the sentence (or the adjacent sentence), but is does not occupy the object position. The missing objects in (41) – (44) can be matched to *us*, *Socrates*/*him*, *me* (the writer) and *Met officers* to whom the written material was addressed to. In (45) the reference is perhaps slightly less explicit, but the warning appears to be directed mostly at the decision makers in Slovenia. In 18 cases of the 113 tokens in the second category the reference was rather clearly mentioned (as in (41) – (44) or can easily be inferred.

In the third and final category I have included the remaining 42 tokens that subtly fall between the two categories introduced above. The reference in these cases is neither generic in the 'wide open' sense as in the first category, nor as restricted as in the second.

(46) Chebrikov was particularly scathing about the 'informal groups' that had been attracting public attention, *warning* that 'extremist elements' had worked their way into the leaderships of some of them and that they were encouraging them to engage in actions 'objectively against the interest of our society.' (FYT 1139)

(47) At a press conference on May 29, the United States Defence Secretary Dick Cheney *warned* that the cost of converting national armies into mobile units could be high. (HL7 2316)

(48) A new report by the US National Research Council, commissioned by the US Congress, has *warned* that children face disproportionate risks from pesticide residues in food. (J3L 11)

(49) On June 5 in Oslo, in a speech accepting the 1990 Nobel Peace Prize, Gorbachev said that "now that perestroika has entered its critical phase, the Soviet Union is entitled to expect large-scale support to assure its success", and *warned* that "if perestroika fails the prospect of entering a new peaceful period in history will vanish". (HL8 123)

Many of the examples in this category come from reports (to do with environment, health and other things of public concern) or public speeches. Are Chebrikov and Cheney in (46) and (47) warning people in general or a more restricted group of people, such as political leaders? It is characteristic of these tokens that both interpretations can be argued. Perhaps, though, the warnings are implicitly directed at a certain group of people on a "you know who you are" basis, but it has been considered wise to present the information so that the public is also informed. In (48), it is wise to let people know about the dangers inherent in pesticide residues, but it is surely not assumed that the general

public will take action. Rather, the warning is directed at those who are in position to react to the situation. In cases such as (49), the public announcement of a more restrictedly directed (to those that large-scale support is expected from) warning is perhaps a deliberate choice of political rhetoric.

Six instances were found in the data that are relevant as regards Bach's Generalization, all of the type *against –ing*-clause:

(50) [Marian Yarrow and Carolyn Waxler at the National Institute of Mental Health,] New York, have *warned* against sampling behaviour as if it were a constant physiological state. (B7L 529)

(51) Tina had *warned* against putting that advertisement in the newsagent's. (EDN 478)

We shall set these aside for now, as they will be discussed in more detail in the following section.

6.2.2 Objectless patterns: object control

A search[11] focusing on object control patterns only (i.e. *–ing*-clauses and *to*-infinitives) returns 88 tokens, out of which 29 are relevant. The distribution of the different patterns is found in the table below:

Pattern	**Number of tokens**	**NF per million**
[– *against –ing*-clause]	27	0.34
[– *of –ing*-clause]	1	0.012
[– *to*-infinitive]	1	0.012
Total	**29**	**0.37**

Table 10. Frequencies of objectless object control patterns.

As might have been expected, the [– *against –ing*-clause] pattern is the most dominant one in the data, but [– *of –ing*-clause] and [– *to*-infinitive] also occur with one token each. The text domain of Informative: World Affairs was again most noticeable, now with 9 tokens, followed by Informative: Social Science (6) and Informative: Arts (4). As for the medium of text, 15 of the 29 relevant tokens were found in periodicals. In sum, the background of the texts where the tokens emerged is similar to that found in the previous section, where we took a more general look at omitting the object.

[11] The whole written part of the BNC with 'Newspapers' once again omitted. Search string: {warn/V} (of|about|against|to) (_V*I | _V*G*)

The semantic categorization also yields similar results. The most common type was the 'least generic reference' (second category) with 20 tokens (example (52)),[12] while four tokens were placed in the first category – 'most generic reference' (53):

(52) The mental scrutiny intensifies: Chris Smith *warns* against letting home-life become too cushy: 'learn how to be loving and gentle at home but flick another switch when playing cricket'. (EB3 1729)

(53) The front panel layout is extremely good: single input, then a small LED which operates with the accompanying gain rotary to *warn* of overdriving the valve preamp stage. (C9J 393)

The remaining five tokens were place in the third category, where the reference in essence is made to a certain group of people, but it is also considered worthwhile to notify 'the public' (see the discussion in 6.2.1.2 for a more detailed discussion on this category as well as the others):

(54) Even those writers who specifically *warned* against using the historical experiences of contemporary rich countries as a guide for the Third World could not resist drawing some conclusions from the realm of foreign trade. (HTV 214)

All in all, then, the semantic categorization of the tokens featuring object control is similar to that seen in the previous section. Furthermore, as the constructions found might lead us to expect, the sense of the verb remains the same throughout, as all examples (perhaps with the exception of example (53) that could be seen as representing Sense I) fall under Sense II. Though the tokens found further corroborate the claim that *warn* is an exception to Bach's Generalization, we are as of now restricted to making observations on semantic features and the context/domain in which tokens in violation of the generalization are found. Rudanko (1989, 129-130) adds to the semantic discussion by noting that *warn*, along with *caution* and *advise* that he mentions as similar exceptions to the generalization, is a non-implicative verb. For example, in *John warned Jane not to intervene,* we do not know whether Jane did eventually intervene or not. This is an interesting point, which helps to semantically tie together at least the three aforementioned verbs that are an exception to Bach's Generalization. That *warn* can be grouped with *advise* and *caution* in this

[12] Cf. the discussion with examples (38)–(45) in 6.2.1.2. The 'least generic reference' category features tokens where the reference can be deduced from the context, as in (38)–(40), or tokens where the referent is more explicitly mentioned elsewhere in the sentence, as in (41)–(44). Example (52) here is reminiscent of examples (38)–(40).

respect is in keeping with the discovery that all tokens (with the exception of (53), perhaps) feature Sense II, which is the 'advise' sense.

That the [– *against –ing*-clause] patterns is found in all but two tokens is interesting in that the *against –ing*-clause construction, be it with or without the object NP, is rare with *warn* compared to *to*-infinitives (also exclusively found with the 'advise' sense), which – as we have seen in this study – have been abundant for centuries, albeit their percentage has decreased. Below is the only token found with a *to*-infinitive:

(55) Thomas Szasz *warns* to beware of the psychiatrist who analyses jokes instead of laughing at them. (FT0 1661)

In addition to mapping out the semantic characteristics of verbs such as *warn* that seem to go against Bach's Generalization, a possible topic for future study is perhaps to further investigate the characteristics of the preposition *against*, together with the *–ing*-clause and the *to*-infinitive in pursuit of a more profound understanding as to why *warn* and some other verbs are exceptional, and, why is it that – at least with *warn* – this exceptional behaviour manifests itself almost exclusively in one pattern.

7 Discussion of main findings and concluding remarks

I have now analysed three different sections of the CLMET followed by a section of the BNC for contemporary data – amounting to a total of 885 relevant tokens. As an additional point of interest, we have taken a closer look at the objectless usage in Present-day English. We shall now sum up the main findings in order to form a coherent picture of the changes that have taken place.

Perhaps the most apparent development is that the overall frequency of *warn* has increased to a great extent. This is illustrated in Figure 1 below:

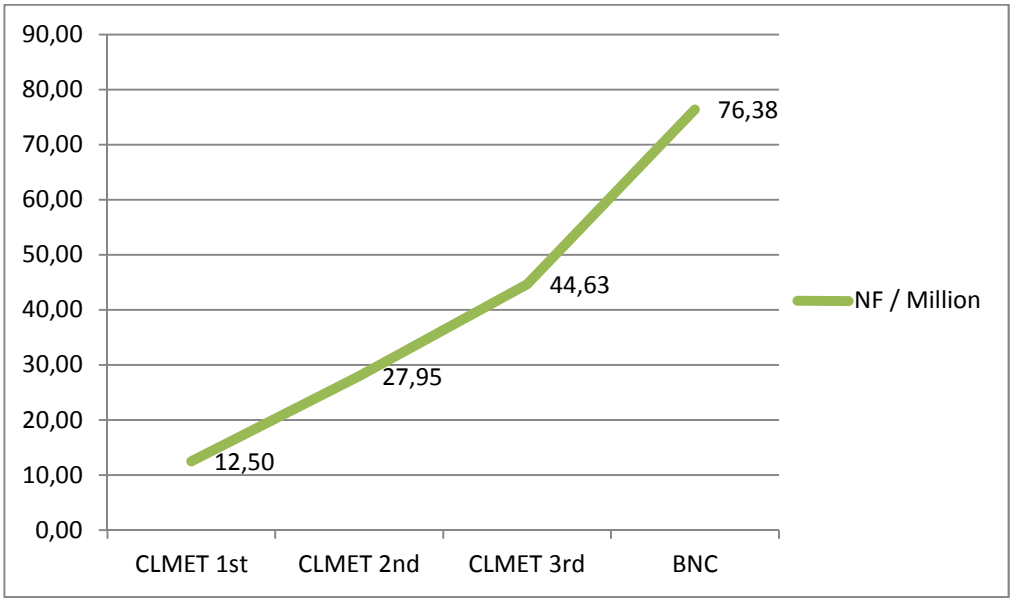

Figure 1. Normalized frequencies of *warn* across the different (sub-)corpora.

As we can see, there is a steady increase in the frequency throughout the time period investigated. The frequencies have gone up from 12.50 in the first part of the CLMET to 27.95 in the second, and on to 44.63 in the third. The third part of the CLMET covers data from the years 1850–1920 and the BNC from 1960–1993. However, as noted earlier, there is a heavy emphasis in the final eight years so that we can say that there exists a gap of some 65 years between the CLMET and the BNC data. This explains why in Figure 1 there appears to be an even steeper incline after the 3rd part of the CLMET. The NF in the BNC is, finally, at 76.38 per million showing that *warn* has become over six times more common since the 18th century.

As regards the patterns, there is a consistent increase in the number of patterns up until the BNC, where a slight drop can be seen: 10, 16 and 19 complement types were found in the 1st, 2nd and 3rd parts of the CLMET, respectively, while in the BNC the corresponding number is 17. This fluctuation can largely be attributed to the more marginal patterns, many of which were – quite understandably – not found in any of the reference works. This applies especially to the multi-prepositional and/or phrasal constructions. Zero complements were found in all sub-corpora, though they have not been counted as complement types per se. In similar vein, comment clauses and the 'speech' pattern were not considered as separate patterns, but are included in zero complements or the simple NP complement. These were both marginal in the CLMET, but suddenly rocketed in the BNC.

As for the distribution of individual constructions, Figure 2 illustrates the developments that have taken place in the five patterns that were the most frequent throughout the data.[13] As the NF of *warn* has increased so consistently, we shall consider the patterns from the point of view of percentage of all tokens in each set – rather than their normalized frequencies. This should bring forth the changes between the patterns more clearly.

[13] The [– *against* NP] pattern was in fact dropped to 6th place in the BNC data in favour of [– NP *about* NP], but its inclusion in Figure 2 is justified as it was among the five top patterns in all three sections of the CLMET.

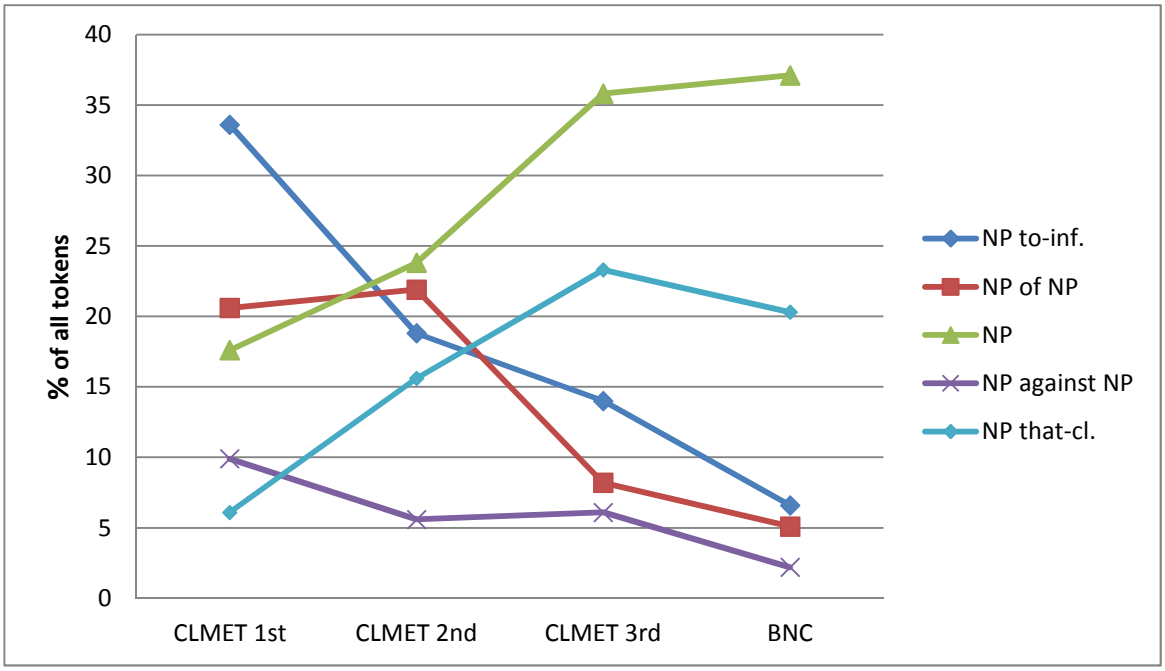

Figure 2. The percentages of the five most frequent complements in the data.

As we can see, considerable change has taken place with the main patterns. The *to*-infinitive started out as the most prominent pattern, but has since come down rapidly. This is rather surprising considering that it is only found with Sense II and can rarely be replaced by another construction (essentially, this is possibly only with the negated *not/never* etc. *to*-infinitives, which can be replaced by *against –ing*-clauses). The change, therefore, cannot be directly linked to the Great Complement Shift.

Another very interesting development is the notable rise of *that*-clauses. Vosberg (2009, 212), among others, has noted that *that*-clauses have been in decline since Middle English, their place since taken by the *to*-infinitive. This does not seem to hold true for *warn*, as quite the opposite has happened: a steady rise throughout the CLMET data is interrupted only as we get to contemporary data. (However, were we to include the tokens of the objectless [– *that*-clause] patterns from the BNC, we would get a percentage similar to that in the 3rd section of the CLMET.)

Though they may not have as much relevance outside of this study as do the developments of the aforementioned sentential patterns, two other (non-sentential) patterns among the top five have also seen clear changes. Firstly, the simple NP complement has increased in percentage and,

starting from the 2nd section of the CLMET, has occupied first place. In the BNC data, nearly a third of the NP tokens feature either comment clauses (21 tokens) or reporting of direct speech (12). As already mentioned, these two types suddenly ballooned to high numbers in the BNC, although they had previously been marginal at best. Especially the usage of the "bare" speech type (zero complement) now appears to be common. Secondly, the NP *of* NP pattern has come down, with the biggest drop between the 2nd and 3rd sections of the CLMET. It is possible that the emergence of the NP *about* NP pattern has brought about at least some of this change, as *about* and *of* can be considered very similar in this context. As for the NP *against* NP pattern, the changes have been less drastic. Nonetheless, it has come down from where it was in the 18th century, and in the BNC it has in fact been superseded by the NP *about* NP pattern.

Turning to the senses, there has been a shift between the two major senses, seen in Figure 3 below.

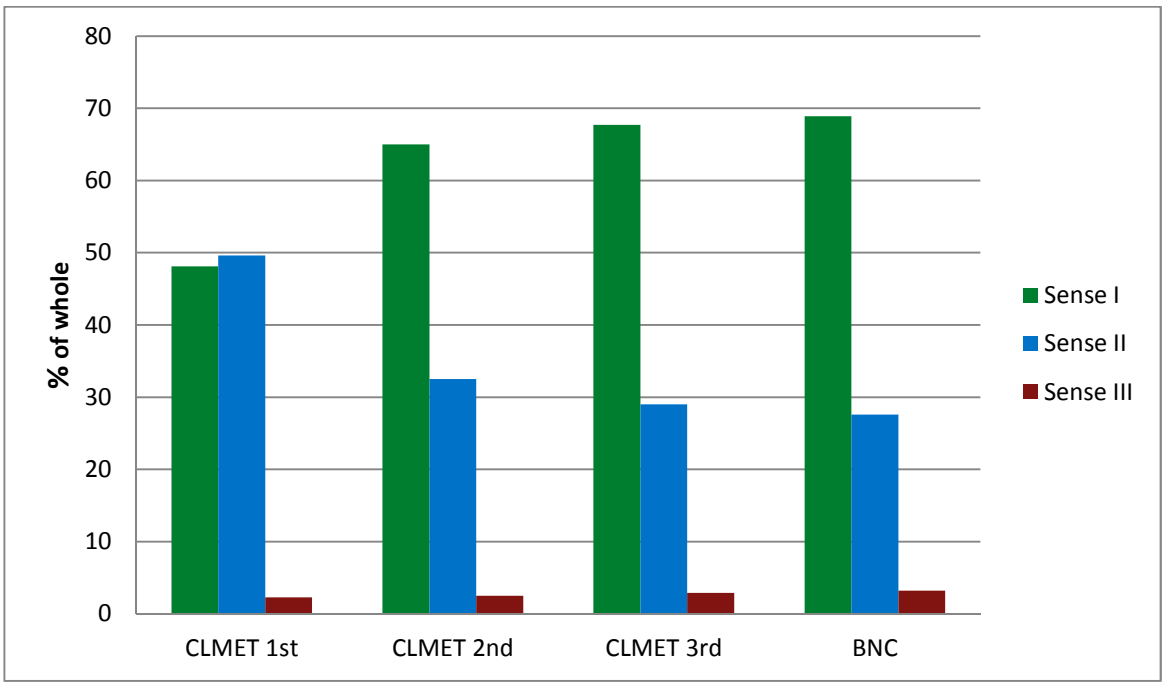

Figure 3. The percentages of the three simplified senses across the data.

It is evident that the change in the patterns is reflected in the distribution of the senses. The single most important factors here are the decrease in the frequency of *to*-infinitives and the increase in

that of the simple NP complement: the *to*-infinitive is found exclusively with Sense II, and most, though by no means all NP complements feature Sense I. With regard to the other major patterns, NP *of* NP is, as we would expect, restricted to Sense I. *That*-clauses and the NP *against* NP pattern are found with both Senses I and II, though the former is clearly more common. Of the less frequent patterns, NP *against –ing*-clause can be considered an equivalent to the negated version of the *to*-infinitive, and is, as a result, only found with Sense II. Sense III is marginal throughout the data, as are the patterns in which it is found. In addition to the three main senses, two tokens of the less likely *OED* senses were found. These were the 'summon' sense in the 3rd part of the CLMET and the 'horse racing' sense in the BNC.

One of the goals of this study was to also test the relevance of some selected theories that bear on complementation. The Complexity Principle stands out as the only one that can be seen to have true value in this particular set of data. No major tendencies were visible in the first two sections of the CLMET, but towards the end of the period under investigation a more consistent correlation is visible: increased complexity – especially insertions – was found with *that*-clauses to an increasing degree. This is especially true with the BNC, where all insertions in the data were found with *that*-clauses. Extractions and *horror aequi* did not play a major role, and there was an insufficient number of tokens to draw any conclusions with regard the Great Complement Shift and the possible rivalry between negated *to*-infinitives and the *against –ing*-clause pattern. However, as already noted, the drop in the frequency of *to*-infinitives and the rise of that in *that*-clauses can be considered surprising in relation to the trends that these two patterns are said to follow, and more light could be shed into the matter by additional research.

The additional searches made on the objectless constructions offer some information on the patterns and their background. In general, omitting the object appears to restrict itself to contexts that are rather formal in nature and often have a connection with politics. *That*-clauses are clearly the one pattern where the omission is most likely. The reference of the omitted, implicit object is

subject to gradience, and more often than not a relatively clear restriction on the referent can be discerned. As for the non-finite sentential complements (which were searched separately), featuring object control and thus relevant to Bach's Generalization, the semantic background is similar. Here, the *against –ing*-clause pattern is found almost exclusively (only one token was found with the *to*-infinitive although it is generally a more frequent pattern with *warn*), which provides a possible topic for further discussion on other semantically similar verbs that are said to be exceptions to the generalization.

The research questions listed at the beginning of this thesis have now been answered. Many interesting developments on the complementation of *warn* have come to light, some of which bear relevance to the study of complementation more broadly and could thus be used as points of interest for future study.

References

Allerton, David. 1988. "'Infinitivitis' in English," in J. Klegraf and D. Niehls, eds., *Essays on the English Language and Applied Linguistics on the Occasion of Gerhard Nickels 60th Birthday*. Heidelberg: Julius Groos.

Ball, C. N. 1994. "Automated Text Analysis: Cautionary Tales." *Literary and Linguistic Computing* 9, 295-302.

Barnhart, Robert K. (ed.). 2006. *Chambers Dictionary of Etymology*. New York: Chambers Harrap.

Biber, Douglas. 2010. "Corpus-based and Corpus-driven Analyses of Language Variation and Use." In *The Oxford Handbook of Linguistic Analysis*, ed. Bernd Heine and Heiko Narrog, 159-191. Oxford: Oxford University Press.

Biber, Douglas, S. Conrad, E. Finegan, S. Johansson and G. Leech. 1999. *Longman Grammar of Spoken and Written English*. London: Longman

Biber, Douglas et al. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Bolinger, Dwight. 1968. "Entailment and the Meaning of Structures." *Glossa* 2, 119-127.

Burnard, Lou (ed.). 2007. *Reference Guide for the British National Corpus (XML edition)*. Oxford: Oxford University Press. [accessed online at: http://www.natcorp.ox.ac.uk/docs/URG/]

Carnie, Andrew. 2002. *Syntax: A Generative Introduction.* Oxford: Blackwell Publishers.

Chomsky, Noam. 1986. *Knowledge of Language*. New York: Praeger.

Chomsky, Noam and Howard Lasnik. 1977. "Filters and Control." *Linguisic Inquiry* 8-3, 425–504.

Conrad, Susan. 2000. "Will Corpus Linguistics Revolutionize Grammar Teaching in the 21st Century?" *TESOL Quarterly* 34, 3: 548-560.

Davies, William and Stanley Dubinsky. 2004. *The Grammar of Raising and Control.* Oxford: Blackwell.

De Smet, Hendrik. 2005. "A Corpus of Late Modern English Texts." *ICAME Journal No*. 29: 69-82.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago and London: The University of Chicago Press.

Haegeman, Liliane. 1991. *Introduction to Government and binding Theory.* Oxford: Basil Blackwell.

Herbst, Thomas et al. 2004. *A Valency Dictionary of English*. Berlin: de Gruyter.

Huang, C.-T. James. 1997. "Introduction to Syntax." LSA Linguistic Institute.

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Hunston, Susan. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Jackson, Howard. 2002. *Lexicography: An Introduction*. London and New York: Routledge.

Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. London: Longman.

Lakoff, George and John Robert Ross. 1966. "Criterion for Verb Phrase Constituency." *Harvard Computation Laboratory Report to the National Science Foundation on Mathematical Linguistics and Automatic Translation* (No. NSF-17). Cambridge: Harvard University Computation Laboratory.

Leech, Geoffrey. 1974. *Semantics.* Harmondsworth: Penguin.

Leech, Geoffrey, M. Hundt, C. Mair and N. Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.

OALD = The Oxford Advanced Learner's Dictionary [accessed online at: http://oald8.oxfordlearnersdictionaries.com/]

OED = The Oxford English Dictionary [accessed online at: http://www.oed.com/]

Perlmutter, David M. 1971. *Deep and Surface Structure Constraints in Syntax.* New York: Holt, Rinehart and Winston.

Postal, Paul. 1970. "On Coreferential Complement Subject Deletion," *Linguistic Inquiry* 1, 439-500.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. 2005. *A Comprehensive Grammar of the English Language.* London: Longman.

Rizzi, Luigi. 1986. "Null Objects in Italian and the Theory of pro." *Linguistic Inquiry* 17, 3: 501-557.

Radford, Andrew. 1988. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.

Rohdenburg, Günter. 1996. "Cognitive complexity and increased grammatical explicitness in English." *Cognitive Linguistics* 7-2, 149-182

Rohdenburg, Günter. 2003. "Cognitive complexity and horror aequi as factors determining the use of interrogative clause linkers in English." In G. Rohdenburg and Britta Mondorf, *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter, 305-328.

Ross, John Robert. 1973. "Nouniness." In Aarts, Denison, Keizer, Popova (eds.), *Fuzzy Grammar*. Oxford: Oxford University Press

Rudanko, Juhani. 1989. *Complementation and Case Grammar*. Albany, N.Y.: SUNY Press.

Somers, Harold. 1984. "On the Validity of the Complement-Adjunct Distinction in Valency Grammar." *Linguistics* 22, 507-530.

Vosberg, Uwe. 2003. “The Role of Extractions and *Horror Aequi* in the Evolution of *-ing* complements in Modern English.” In G. Rohdenburg and Britta Mondorf, *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter, 305-328.

Vosberg, Uwe. 2009. “Non-finite Complements.” In G. Rohdenburg and Julia Schlüter (eds.), *One Language, Two Grammars? Differences between British and American English.* Cambridge: Cambridge University Press.

Corpora used

BNC = The British National Corpus (XML edition). [accessed using the BNCweb interface at http://bncweb.uta.fi]

CLMET = The Corpus of Late Modern English Texts (Extended Version and the CLMET 3.0)